```
**********************************************
                  MPSRCH (TM)
**********************************************

       Release 3.1A John F. Collins, Biocomputing Research Unit.
       Copyright (c) 1993-1998 University of Edinburgh, U.K.
              Distribution rights by Oxford Molecular Ltd

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

   on:       Fri Nov 13 17:13:47 1998;   MasPar time 13.94 Seconds
...lar output not generated.              828.366 Million cell updates/sec

Title:          >US-08-989-362-2
Description:    (1-316) from US08989362.pep
Perfect Score:  2294
Sequence:     1 MRRASRDYGKYLRSSEEMGS.........LLDPDQDATYFGAFKVQDID 316

Scoring table:  PAM 150
                Gap 11

Searched:       120441 seqs, 36531193 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:       pir56
                1:pir1 2:pir2 3:pir3 4:pir4 5:nrl3d

Statistics:     Mean 46.143;  Variance 97.466;  scale 0.473

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                                   SUMMARIES
```

Result No.	Score	Query Match %	Length	DB	ID	Description	Pred. No.
1	174	7.6	278	2	A49266	fas ligand - rat	7.29e-11
2	175	7.6	281	2	JC2340	Fas ligand - human	5.11e-11
3	172	7.5	279	2	A53062	Fas ligand - mouse	1.48e-10
4	165	7.2	146	5	1ALY	cd40 ligand extracell	1.74e-09
5	165	7.2	261	2	I53476	CD40 ligand - human	1.74e-09
6	162	7.1	261	2	S53090	CD40 ligand - bovine	4.94e-09
7	160	7.0	235	1	QWMSN	tumor necrosis factor	9.87e-09
8	152	6.6	234	2	A25451	tumor necrosis factor	1.52e-07
9	150	6.5	235	2	JU0029	CD40 ligand - mouse	3.00e-07
10	150	6.5	260	2	S21738	CD40 ligand - mouse	3.00e-07
11	145	6.3	235	2	I54490	tumor necrosis factor	1.60e-06
12	143	6.2	234	2	JQ1344	tumor necrosis factor	3.10e-06
13	141	6.1	146	5	1CDAB	Cd40 ligand complex w	6.00e-06
14	141	6.1	146	5	1CDAC	Cd40 ligand complex w	6.00e-06
15	141	6.1	146	5	1CDAA	Cd40 ligand complex w	6.00e-06
16	136	5.9	306	2	I49139	lymphotoxin-beta - mo	3.06e-05
17	134	5.8	193	2	S06192	tumor necrosis factor	5.84e-05
18	130	5.7	151	5	2TUNA	Tumor necrosis factor	2.09e-04
19	130	5.7	151	5	2TUNE	Tumor necrosis factor	2.09e-04
20	130	5.7	151	5	2TUNF	Tumor necrosis factor	2.09e-04
21	130	5.7	151	5	2TUNC	Tumor necrosis factor	2.09e-04
22	130	5.7	151	5	2TUND	Tumor necrosis factor	2.09e-04
23	130	5.7	151	5	2TUNB	Tumor necrosis factor	2.09e-04
24	128	5.6	233	2	S11688	tumor necrosis factor	3.93e-04
25	126	5.5	233	1	QWHUN	tumor necrosis factor	7.35e-04
26	123	5.4	152	5	1TNFB	tumor necrosis factor	1.86e-03
27	123	5.4	152	5	1TNFC	tumor necrosis factor	1.86e-03
28	123	5.4	152	5	1TNFA	tumor necrosis factor	1.86e-03
29	123	5.4	185	2	S52715	tumor necrosis factor	1.86e-03
30	124	5.4	233	2	S13114	tumor necrosis factor	1.37e-03
31	123	5.4	233	2	S24642	tumor necrosis factor	1.86e-03
32	123	5.4	233	2	S22052	tumor necrosis factor	1.86e-03
33	125	5.3	234	2	JH0529	tumor necrosis factor	1.00e-03
34	122	5.3	144	5	1TNRA	tumor necrosis factor	2.54e-03
35	122	5.3	205	1	QWHUX	lymphotoxin alpha pre	2.54e-03
36	114	5.2	232	2	S12606	tumor necrosis factor	2.83e-02
37	111	4.8	244	2	A46066	lymphotoxin beta - hu	6.81e-02
38	107	4.7	202	2	B27303	tumor necrosis factor	2.15e-01
39	106	4.6	345	2	S12788	transcription factor	2.86e-01
40	106	4.6	963	2	S45167	chitin synthase (EC 2	2.86e-01
41	103	4.5	197	2	JH0309	tumor necrosis factor	6.63e-01
42	104	4.5	202	2	JN0869	tumor necrosis factor	5.02e-01
43	104	4.5	1091	2	S33596	protein-tyrosine kina	5.02e-01
44	100	4.4	264	2	A30582	MHC class II histocom	1.51e+00
45	100	4.4	473	2	A54691	octamer-binding prote	1.51e+00

```
                                   ALIGNMENTS

RESULT    1
ENTRY      A49266      #type complete
TITLE      fas ligand - rat
ORGANISM   #formal_name Rattus norvegicus #common_name Norway rat
DATE       13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change
           10-Sep-1997
ACCESSIONS A49266
REFERENCE  A49266
  #authors Suda, T.; Takahashi, T.; Golstein, P.; Nagata, S.
  #journal Cell (1993) 75:1169-1178
  #title   Molecular cloning and expression of the Fas ligand, a novel
           member of the tumor necrosis factor family.
  #accession A49266
  ##status   preliminary
  ##molecule_type mRNA
  ##residues 1-278 ##label SUD
  ##cross-references GB:U03470; NID:g440178; PID:g440179
KEYWORDS   glycoprotein; transmembrane protein
SUMMARY    #length 278 #molecular-weight 31140 #checksum 9586

Query Match 7.6%;  Best Local Similarity 26.7%;  Score 174;  DB 2;  Length 278;
Matches 36;  Conservative 38;  Mismatches 53;  Indels 8;  Gaps 6;

Db  152 SRSIPLE-WEDTYGTALISGVKYKKGGLVINEAGLYFVYSKVYFRGQ-SCNSQP---LSH 206
        |:|:||| | | ||: |::|:| : :|:|::||:|::|::|:||::|| |:|||
Qy  178 SHKVTLSSWTHDRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYIQL 237

Db  207 KVYM-R-NFKYPGDLVLMEE-KKLNYCTTGQIWAHSSYLGAVFNLTVADHLYVNISQLSL 263
        :|:|: | |: | || | |:| :: :: ::| |
Qy  238 MVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSL 297

Db  264 INFEESKTFFGLYKL 278
        ::| :| || :|: |:| :|:
Qy  298 LDPDQDATYFGAFKV 312

RESULT    2
ENTRY      JC2340      #type complete
TITLE      Fas ligand - human
ORGANISM   #formal_name Homo sapiens #common_name man
DATE       20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
           31-Oct-1997
ACCESSIONS JC2340; S57565; I38554
REFERENCE  JC2340
  #authors Mita, E.; Hayashi, N.; Iio, S.; Takehara, T.; Hijioka, T.;
           Kasahara, A.; Fusamoto, H.; Kamada, T.
```

```
#journal    Biochem. Biophys. Res. Commun. (1994) 204:468-474
#title      Role of Fas ligand in apoptosis induced by hepatitis C virus
            infection.
#accession  JC2340
##molecule_type DNA
##residues  1-281 ##label MIT
##cross-references GB:D38122; DDBJ:D29820; NID:g601892; PID:d1007898;
            PID:g1369902
REFERENCE   S57565
#authors    Schatzlein, C.E.
#submission submitted to the EMBL Data Library, June 1995
#accession  S57565
##status    preliminary
##molecule_type mRNA
##residues  1-281 ##label SCH
##cross-references EMBL:X89102; NID:g887455; PID:g887456
REFERENCE   I38554
#authors    Alderson, M.R.; Tough, T.W.; Davis-Smith, T.; Braddy, S.;
            Falk, B.; Schooley, K.A.; Goodwin, R.G.; Smith, C.A.;
            Ramsdell, F.; Lynch, D.H
#journal    J. Exp. Med. (1995) 181:71-77
#title      Fas ligand mediates activation-induced cell death in human T
            lymphocytes.
#cross-references MUID:95105731
#accession  I38554
##status    preliminary; translated from GB/EMBL/DDBJ
##molecule_type mRNA
##residues  1-281 ##label RES
##cross-references EMBL:U08137; NID:g624627; PID:g624628
GENETICS
#introns    151/1; 116/3
KEYWORDS    glycoprotein; transmembrane protein
FEATURE
    80-102          #domain transmembrane #status predicted #label TMM\
                    predicted
    76,184,250,260  #binding_site carbohydrate (Asn) (covalent) #status
SUMMARY     #length 281 #molecular-weight 31485 #checksum 3826

Query Match    7.6%;    Score 175;    DB 2;    Length 281;
Best Local Similarity 23.9%;  Pred. No. 5.11e-11;
Matches 44;   Conservative 55;   Mismatches 75;   Indels 10;   Gaps 8;

Db   107  LQKELAELR-ESTSOMHTASSLE-KQIGHPSPPPEKKELRKVAHLTGKSNSRSMPLEWED 164
          ||  ::  :::   :|:|   ||::    |   ||:: ::|||:  ::|||:     |
Qy   130  LQHIVGPQRFSGAPAMMEGSWLDVAQRGKPEAQPFAHLTINAASIPSGSHKVTLS-SWYH 188
Db   165  TYGIVLLSGVKYKKGGLVINETGLYFVYSKVYFRGQ-SCNNLP---LSHKVYM-RNS-KY 218
          |:|: ||:| |   ||||||||||||||||||||| :::::   |: ||::  ::|  |
Qy   189  DRGWAKISNMTLSNGKLRVNQDGFYLYANICFRHHETSGSVPTDYLQLMVYVVKTSIKI 248
Db   219  PQDL-VNMEGKMMSYCTTGQMWARSSYLGAVFNLTSADHLYVNVSELSLVNFEESQTFFG 277
          |  | :::| :  :    : ||::|||||||||||||||||:  ::|  ::|   :|||
Qy   249  PSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDPDQDATYFG 308
Db   278  LYKL 281
          :|:
Qy   309  AFKV 312

RESULT 3
ENTRY       A53062    #type complete
TITLE       Fas ligand - mouse
ORGANISM    #formal_name Mus musculus #common_name house mouse
DATE        06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
            10-Sep-1997
ACCESSIONS  A53062
REFERENCE   A53062
#authors    Takahashi, T.; Tanaka, M.; Brannan, C.I.; Jenkins, N.A.;
            Copeland, N.G.; Suda, T.; Nagata, S.
#journal    Cell (1994) 76:969-976
#title      Generalized lymphoproliferative disease in mice, caused by a
            point mutation in the Fas ligand.
#accession  A53062
```

```
##status    preliminary
##molecule_type mRNA
##residues  1-279 ##label TAK
##cross-references GB:U06948; NID:g473564; PID:g473565
SUMMARY     #length 279 #molecular-weight 31442 #checksum 3291

Query Match    7.5%;    Score 172;    DB 2;    Length 279;
Best Local Similarity 25.9%;  Pred. No. 1.48e-10;
Matches 35;   Conservative 39;   Mismatches 53;   Indels 8;   Gaps 7;

Db   153  SRSIPLE-WEDTYGTALISGVKYKKGGLVINETGLYFVYSKVYFRGQ-SCNNQPLN--H- 207
          |  : |:   : ||| |:::| |   |||||||||||||||||||| :  :|  :  :
Qy   178  SHKVTLSSWYHDRGWAKISNMTLSNGKLRVNQDGFYLYANICFRHHETSGSVPTDYLQL 237
Db   208  KVYM-RNS-KYPEDLVLMEEKRL-NYCTTGQIWAHSSYLGAVFNLTSADHLYVNISQLSL 264
          :|:  ::|  |  :: ::::::  : |||| :||:|||||||||||||||||||:  ||
Qy   238  MVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSL 297
Db   265  INFEESKTFFGLYKL 279
          :| :|  |||:|:
Qy   298  LDPDQDATYFGAFKV 312

RESULT 4
ENTRY       1ALY    #type complete
TITLE       cd40 ligand extracellular domain fragment - human
PDB_TITLE   crystal structure of human cd40 ligand
ORGANISM    #formal_name Homo sapiens #common_name man
#note       blood, expressed in Pichia pastoris
REFERENCE   A68008
#authors    Karpusas, M.; Hsu, Y.M.; Thomas, D.
#submission submitted to the Brookhaven Protein Data Bank, June 1997
#cross-references PDB:1ALY
REFERENCE   TN000391
#authors    Karpusas, M.; Hsu, Y.M.; Wang, J.H.; Thompson, J.; Lederman,
            S.; Chess, L.; Thomas, D.
#journal    Structure (London) (1995) 3:1031
#title      2 a crystal structure of an extracellular fragment of human
            cd40 ligand.
REFERENCE   TN000392
#authors    Karpusas, M.; Hsu, Y.M.; Wang, J.H.; Thompson, J.; Lederman,
            S.; Chess, L.; Thomas, D.
#journal    Structure (London) (1995) 3:1426
#title      Erratum. 2 a crystal structure of an extracellular fragment
            of human cd40 ligand.
COMMENT     Resolution: 2.0 angstroms
COMMENT     Determination: x-ray diffraction
COMMENT     R-value: 0.223
KEYWORDS    cd40 ligand; cd401; cytokine; hyper-igm syndrome; tnf
FEATURE
    129-131         #region helix (right hand 3-10)\
    8-12,140-145,52-64, :
    104-116         #region beta sheet\
    38-41,45-48     #region beta sheet\
    121-126,73-81,  :
    88-96           #region beta sheet
SUMMARY     #length 146 #molecular-weight 15797 #checksum 6328

Query Match    7.2%;    Score 165;    DB 5;    Length 146;
Best Local Similarity 29.5%;  Pred. No. 1.74e-09;
Matches 44;   Conservative 47;   Mismatches 44;   Indels 14;   Gaps 12;

Db   7    IAAHVISEASSKTTSVLQW-AEKGYYTMSNNLVTLENGKQLTVKRQGLYIYAQVTFCSN 65
          |.|.:: ::: ||:  :  | ||:|   :: :.||:|.|  | |.|.||:|.:| :|.:
Qy   169  INAASIPSGSHKVT-LSSWYHDRGWAKISN-M-TLSNGK-LRVNQDGFYLYANICF-RH 223
Db   66   REASSQAP--FIASL-CL-KSPGRF-ERILLRAANTHS-SAKPCGQ-QSIHLGGVFELQ 117
          .|.||. |  |:.||  | |||    ||:|| ..:||: ||||  | |||||||||||:
Qy   224  HETSGSVPTDYLQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLR 283
Db   118  PGASVFNVTDPSQVSHGTGFTSFGLLKL 146
          .|.||:.||.|.|:|     |.||::|:
Qy   284  AGEEISIQVSNPSLLDPDQDATYFGAFKV 312
```

```
                                      #map_position Xq26-Xq26
KEYWORDS     glycoprotein; transmembrane protein
FEATURE
     13-44          #domain transmembrane #status predicted #label TMM\
     45-261         #domain extracellular #status predicted #label EXT\
     6,240          #binding_site carbohydrate (Asn) (covalent) #status
                        predicted
SUMMARY      #length 261 #molecular-weight 29273 #checksum 8322

Query Match               7.2%;  Score 165;  DB 2;  Length 261;
Best Local Similarity    29.5%;  Pred. No. 1.74e-09;
Matches 44; Conservative 47; Mismatches 44; Indels 14;  Gaps 12;

Db  122  IAAHVISEASSKTTSVLQW-AEKGYYTMSNNLVTLENGKQLTVKRQGLYYIYAQVTFCSN 180
         | | |:|||:| :: | :| :| :|| ||| ||::||:||::|| :||::||  :||
Qy  169  INAASIPSGSHKVT-LSSWYHDRGWAKISN-M-TLSNGK-LRVNQDGFYYLYANICF-RH 223

Db  181  REASSQAP--FIASL-CL-KSPGRF--ERILLRAANTHS-SAKPCGQ-QSIHLGGVFELQ 232
         :|:| |    :  |  ::|:  :    |  :::|:| | :   ::|| |   :||:|| |
Qy  224  HETGGSVPTDYLQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLR 283

Db  233  PGASVFVNVTDPSQVSHGTGFTSFGLLKL 261
         :  |||:|  |||| |||||:|| |:| :|
Qy  284  AGEEISIQVSNPSLLDPDQDATYFGAFKV 312

RESULT    6
ENTRY     S53090         #type complete
TITLE     CD40 ligand - bovine
ORGANISM  #formal_name Bos primigenius taurus #common_name cattle
DATE      08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change
              10-Sep-1997
ACCESSIONS  S53090
REFERENCE   S53090
  #authors      Mertens, B.E.L.C.; Muriuki, M.
  #submission   submitted to the EMBL Data Library, February 1995
  #description  Cloning of bovine CD40L and homology to bovine TNFA and TNFB.
  #accession    S53090
  ##status          preliminary
  ##molecule_type   mRNA
  ##residues        1-261 ##label MER
  ##cross-references EMBL:Z48465; NID:g732569; PID:g732570
SUMMARY      #length 261 #molecular-weight 29242 #checksum 8929

Query Match               7.1%;  Score 162;  DB 2;  Length 261;
Best Local Similarity    30.2%;  Pred. No. 4.94e-09;
Matches 45; Conservative 40; Mismatches 50; Indels 14;  Gaps 13;

Db  122  IAAHVISEASSKTTSVLQW-APKGYYTLSNNLVTLENGKQLAVKRQGFYYIYTQVTFCSN 180
         | | |:|||:| :: | :| :| | ||  || ||::||:|| |:||::||  :||
Qy  169  INAASIPSGSHKVT-LSSWYHDRGWAKISN-M-TLSNGK-LRVNQDGFYYLYANICF-RH 223

Db  181  RETLSQAP--FIASL-CL-KSP-S-GSERILRAANTHSSSKPCG-Q-QSIHLGGVFELQ 232
         :|:| |    :  |  ::|:  :  |  :::|:| | :::|| |   :||:|| |
Qy  224  HETGGSVPTDYLQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLR 283

Db  233  SGASVFVNVTDPSQVSHGTGFTSFGLLKL 261
         :  |||:|  |||| |||||:|| |:| :|
Qy  284  AGEEISIQVSNPSLLDPDQDATYFGAFKV 312

RESULT    7
ENTRY     QWMSN          #type complete
TITLE     tumor necrosis factor alpha precursor - mouse
ALTERNATE_NAMES   cachectin; TNF alpha
ORGANISM  #formal_name Mus musculus #common_name house mouse
DATE      31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change
              20-Mar-1998
ACCESSIONS  A22908; S03791; A27303; A25164; A23127; A34251; I59058;
            A36696
REFERENCE   A22908
  #authors   Shirai, T.; Shimizu, N.; Shiojiri, S.; Horiguchi, S.; Ito, H.
  #journal   DNA (1988) 7:193-201
```

```
RESULT    5
ENTRY     I53476         #type complete
TITLE     CD40 ligand - human
ALTERNATE_NAMES   glycoprotein 39; hCD40-L protein; T-cell antigen gp39; TRAP
                      protein
ORGANISM  #formal_name Homo sapiens #common_name man
DATE      02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
              13-Mar-1998
ACCESSIONS  I53476; S28017; JH0793; S26694; S28852; S25684; S30593
REFERENCE   I53476
  #authors   Gauchat, J.
  #journal   FEBS Lett. (1992) 315:259-266
  #title     Human CD40-ligand: Molecular cloning, cellular distribution
                 and regulation of expression by factors controlling IgE
                 production.
  #cross-references MUID:93138085
  #accession I53476
  ##status          preliminary; translated from GB/EMBL/DDBJ
  ##molecule_type   mRNA
  ##residues        1-261 ##label RES
  ##cross-references GB:L07414; NID:g180123; PID:g180124
REFERENCE   S28017
  #authors   Hollenbaugh, D.; Grosmaire, L.S.; Kullas, C.D.; Chalupny,
                 N.J.; Braesch-Andersen, S.; Noelle, R.J.; Stamenkovic, I.;
                 Ledbetter, J.A.; Aruffo, A.
  #journal   EMBO J. (1992) 11:4313-4321
  #title     The human T cell antigen gp39, a member of the TNF gene
                 family, is a ligand for the CD40 receptor: expression of a
                 soluble form of gp39 with B cell co-stimulatory activity.
  #accession S28017
  ##molecule_type   mRNA
  ##residues        1-261 ##label HOL
  ##cross-references EMBL:Z15017; NID:g38483; PID:g38484
REFERENCE   JH0793
  #authors   Spriggs, M.K.; Armitage, R.J.; Strockbine, L.; Clifford,
                 K.N.; Macduff, B.M.; Sato, T.A.; Maliszewski, C.R.;
                 Fanslow, W.C.
  #journal   J. Exp. Med. (1992) 176:1543-1550
  #title     Recombinant human CD40 ligand stimulates B cell proliferation
                 and immunoglobulin E secretion.
  #accession JH0793
  ##molecule_type   mRNA
  ##residues        1-261 ##label SPR
  ##cross-references GB:X67878; NID:g38411; PID:g38412
  ##experimental_source peripheral blood T-cell
REFERENCE   S26694
  #authors   Graf, D.; Korthaeuer, U.; Mages, H.W.; Senger, G.; Kroczek,
                 R.A.
  #journal   Eur. J. Immunol. (1992) 22:3191-3194
  #title     Cloning of TRAP, a ligand for CD40 on human T cells.
  #accession S26694
  ##status          preliminary
  ##molecule_type   mRNA
  ##residues        1-261 ##label GRA
  ##cross-references EMBL:X68550; NID:g37269; PID:g37270
REFERENCE   S28852
  #authors   Gauchat, J.F.; Aubry, J.P.; Mazzei, G.; Life, P.; Jomotte,
                 T.; Elson, G.; Bonnefoy, J.Y.
  #journal   FEBS Lett. (1993) 315:259-266
  #title     Human CD40-ligand: molecular cloning, cellular distribution
                 and regulation of expression by factors controlling IgE
                 production.
  #accession S28852
  ##molecule_type   mRNA
  ##residues        1-261 ##label GAU
  ##cross-references EMBL:L07414;
  ##note      the sequence from Fig. 3 is inconsistent with that from
                 Fig. 2 in having 6-Gln
GENETICS
  #gene      GDB:CD40LG; HIGM1; IMD3
  #cross-references GDB:120632; OMIM:308230
```

#title   Cloning and expression in Escherichia coli of the gene for
          mouse tumor necrosis factor.
#cross-references MUID:88224564
#accession A22908
##molecule_type DNA
##residues 1-235 ##label SHI
##cross-references GB:M20155
REFERENCE S03791
#authors Shakhov, A.N.; Nedospasov, S.A.
#journal Bioorg. Khim. (1987) 13:701-705
#title Molecular cloning of the genes coding for tumor necrosis
       factors: complete nucleotide sequence of the genomic copy
       of TNF-alpha in mice.
#cross-references MUID:87298639
#accession S03791
##molecule_type DNA
##residues 1-235 ##label SHA
##cross-references GB:M38296; NID:g202086; PID:g202087
##note article in Russian with English abstract
REFERENCE A93679
#authors Semon, D.; Kawashima, E.; Jongeneel, C.V.; Shakhov, A.N.;
         Nedospasov, S.A.
#journal Nucleic Acids Res. (1987) 15:9083-9084
#title Nucleotide sequence of the murine TNF locus, including the
       TNF-alpha-(tumor necrosis factor) and TNF-beta-
       (lymphotoxin) genes.
#cross-references MUID:88067722
#accession A27303
##molecule_type DNA
##residues 1-235 ##label SEM
REFERENCE A25164
#authors Pennica, D.; Hayflick, J.S.; Bringman, T.S.; Palladino, M.A.;
         Goeddel, D.V.
#journal Proc. Natl. Acad. Sci. U.S.A. (1985) 82:6060-6064
#title Cloning and expression in Escherichia coli of the cDNA for
       murine tumor necrosis factor.
#cross-references MUID:85298296
#accession A25164
##molecule_type mRNA
##residues 1-235 ##label PEN
##cross-references GB:M11731; NID:g202084; PID:g202085
REFERENCE A23127
#authors Fransen, L.; Muller, R.; Marmenout, A.; Tavernier, J.; van
         der Heyden, J.; Kawashima, E.; Chollet, A.; Tizard, R.; van
         Heuverswyn, H.; van Vliet, A.; Ruysschaert, M.R.; Fiers, W.
#journal Nucleic Acids Res. (1985) 13:4417-4429
#title Molecular cloning of mouse tumour necrosis factor cDNA and
       its eukaryotic expression.
#cross-references MUID:85242112
#accession A23127
##molecule_type mRNA
##residues 1-235 ##label FRA
##cross-references GB:X02611; NID:g54844; PID:g54845
REFERENCE A34251
#authors Cseh, K.; Beutler, B.
#journal J. Biol. Chem. (1989) 264:16256-16260
#title Alternative cleavage of the cachectin/tumor necrosis factor
       propeptide results in a larger, inactive form of secreted
       protein.
#cross-references MUID:89380231
#accession A34251
##molecule_type protein
##residues 70-87 ##label CSE
REFERENCE I59058
#authors Caput, D.; Beutler, B.; Hartog, K.; Thayer, R.; Brown-Shimer,
         S.L.; Cerami, A.
#journal Proc. Natl. Acad. Sci. U.S.A. (1986) 83:1670-1674
#title Identification of a common nucleotide sequence in the
       3'-untranslated region of mRNA molecules specifying
       inflammatory mediators.
#cross-references MUID:86149365
#accession I59058
##status preliminary; translated from GB/EMBL/DDBJ

##molecule_type mRNA
##residues 1-230,'R',232-235 #label RES
##cross-references GB:M13049; NID:g202082; PID:g202083
REFERENCE A36696
#authors Sherry, B.; Jue, D.M.; Zentella, A.; Cerami, A.
#journal Biochem. Biophys. Res. Commun. (1990) 173:1072-1078
#title Characterization of high molecular weight glycosylated forms
       of murine tumor necrosis factor.
#cross-references MUID:91097531
#accession A36696
##molecule_type protein
##residues 80-85,'X',87-99 ##label SHE
GENETICS
#introns 62/3; 81/1; 97/1
#note the first intron occurs in the 5'-untranslated region
CLASSIFICATION #superfamily tumor necrosis factor
KEYWORDS cytokine; cytotoxin; glycoprotein; lymphokine; macrophage;
         membrane protein
FEATURE
80-235   #product tumor necrosis factor #status experimental
                #label MAT\
86   #binding_site carbohydrate (Asn) (covalent) #status
                predicted\
148-179   #disulfide_bonds #status predicted
SUMMARY #length 235 #molecular-weight 25895 #checksum 8149

Query Match   7.0%; Score 160; DB 1; Length 235;
Best Local Similarity 25.6%; Pred. No. 9.87e-09;
Matches 33; Conservative 34; Mismatches 56; Indels 6; Gaps 6;

Db 107 WLSQRANALLANGMDLKDNQLVVPADGLYLVYSQVLFKGQGCPDYVLLTH-TVSRFAISY 165
           : :|: || : | | : | | ||: ||: : | :: :
Qy 186 WYHDRGWAKISN-MTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYVVKT 244
Db 166 QEKV-NLLSAVKSPCPKDTPEGAELKPWYEPIYLGGVFQLEKGDQLSAEVNLPKYLDFAE 224
          : : :: :|: | || :| : :: | :   :|| : || | = |: :: ::
Qy 245 SIKIPSSHNLMKGGSTKNWSGNSEFH-FY-SINVGGFFKLRAGEEISIQVSNPSLLD-PD 301
Db 225 SGQVFPGVI 233
          ||| |
Qy 302 QDATYFGAF 310

RESULT 8
ENTRY   A25451   #type complete
TITLE   tumor necrosis factor alpha precursor - rabbit
ALTERNATE_NAMES cachectin; TNF alpha
ORGANISM   #formal_name Oryctolagus cuniculus #common_name domestic
                rabbit
DATE   05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change
                20-Mar-1998
ACCESSIONS   A25454; A25451; JS0727
REFERENCE A25454
#authors Ito, H.; Yamamoto, S.; Kuroda, S.; Sakamoto, H.; Kajihara,
         J.; Kiyota, T.; Hayashi, H.; Kato, M.; Seko, M.
#journal DNA (1986) 5:149-156
#title Molecular cloning and expression in Escherichia coli of the
       cDNA coding for rabbit tumor necrosis factor.
#cross-references MUID:86219711
#accession A25454
##molecule_type mRNA
##residues 1-234 ##label ITO
##cross-references GB:M12845; NID:g165759; PID:g165760
REFERENCE A25451
#authors Ito, H.; Shirai, T.; Yamamoto, S.; Akira, M.; Kawahara, S.;
         Todd, C.W.; Wallace, R.B.
#journal DNA (1986) 5:157-165
#title Molecular cloning of the gene encoding rabbit tumor necrosis
       factor.
#cross-references MUID:86219712
#accession A25451
##molecule_type DNA
##residues 1-234 ##label IT2

    ##note   this sequence differs from that shown in having a Gln
              inserted between residues 62 and 63
REFERENCE   JH0309
    #authors Shakhov, A.N.; Kuprash, D.V.; Azizov, M.M.; Jongeneel, C.V.;
              Nedospasov, S.A.
    #journal Gene (1990) 95:215-221
    #title   Structural analysis of the rabbit TNF locus, containing the
              genes encoding TNF-beta (lymphotoxin) and TNF-alpha (tumor
              necrosis factor)
    #cross-references MUID:91065534
    #accession JS0727
    ##status nucleic acid sequence not shown; translation not shown
    ##molecule_type DNA
    ##residues 1-62,'Q',63-234 ##label SHA
    ##cross-references GB:M60340; GB:M35326; NID:g165754; PID:g165756
GENETICS
    #introns 62/3; 80/1; 96/1
CLASSIFICATION #superfamily tumor necrosis factor
KEYWORDS   cytokine; cytotoxin; glycoprotein; lymphokine; macrophage;
            membrane protein
FEATURE
    1-81   #domain signal sequence #status predicted #label SIG\
    82-234   #product tumor necrosis factor #status predicted #label
              MAT
SUMMARY   #length 234 #molecular-weight 25688 #checksum 664

Query Match   6.6%; Score 152; DB 2; Length 234;
Best Local Similarity 26.0%; Pred. No. 1.52e-07;
Matches   33; Conservative   33; Mismatches   55; Indels   6; Gaps   6;

Db 106 WLSQRANALLANGMKLTDNQLVVPADGLYLIYSQVLFSGQGCRSVVLLTH-TVSRFAVSY 164
        | :|: | :::|| ::| | : :||: |||| : :: |:: :::|:
Qy 186 WYHDRGWAKISN-MTLSNGKLRVNQDGFYLYANICFRHHETSGSVPTDYIQLMYVVKT 244

Db 165 PNKV-NLLSAIKSPCHRETPEEAEPMAWYEPIYLGGVFQLEKGDRLSTEVNQPEYLDLAE 223
        : : :|| :|| | |:: :| : | : ||| || : |||
Qy 245 SIKIPSSHNLMKGGSTKNWSGNSE-FHFY-SINVGGFFKLRAGEEISIQVSNPSLLD-PD 301

Db 224 SGQVYFG 230
        |||
Qy 302 QDATYFG 308

RESULT   9
ENTRY   JU0029   #type complete
TITLE   tumor necrosis factor alpha precursor - rat
ALTERNATE_NAMES cachectin; TNF alpha
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
         07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change
         20-Mar-1998
ACCESSIONS JU0029; JN0868; S21674
REFERENCE JU0029
    #authors Shirai, T.; Shimizu, N.; Horiguchi, S.; Ito, H.
    #journal Agric. Biol. Chem. (1989) 53:1733-1736
    #title   Cloning and expression in Escherichia coli of the gene for
              rat tumor necrosis factor.
    #accession JU0029
    ##molecule_type DNA
    ##residues 1-235 ##label SHI
REFERENCE JN0868
    #authors Kwon, J.; Chung, I.Y.; Benveniste, E.N.
    #journal Gene (1993) 132:227-236
    #title   Cloning and sequence analysis of the rat tumor necrosis
              factor-encoding genes.
    #accession JN0868
    ##molecule_type DNA
    ##residues 1-235 ##label KWO
    ##cross-references GB:L00981; NID:g205253; PID:g205254
REFERENCE S21674
    #authors Essler, H.C.; Grewe, M.; Gaussling, R.; Pavlovic, M.; Decker,
              K.
    #journal Biol. Chem. Hoppe-Seyler (1992) 373:271-281
    #title   Rat tumor necrosis factor-alpha. Transcription in rat Kupffer

                  cells and in vitro posttranslational processing based on a
                  PCR-derived cDNA.
    #cross-references MUID:92329007
    #accession S21674
    ##molecule_type mRNA
    ##residues   1-38,'P',40-162,'T',164-201,'S',203-235 ##label EST
    ##cross-references GB:X66539; GB:S40199; NID:g395369; PID:g395370
COMMENT   Tumor necrosis factor is secreted by macrophages in response to
          endotoxin and produces hemorrhagic necrosis of tumors.
GENETICS
    #gene   TNF-alpha
    #introns 62/3; 81/1; 97/1
CLASSIFICATION #superfamily tumor necrosis factor
KEYWORDS   cytokine; cytotoxin; glycoprotein; lymphokine; macrophage;
            membrane protein
FEATURE
    80-235   #product tumor necrosis factor #status predicted #label
              MAT\
    86   #binding_site carbohydrate (Asn) (covalent) #status
              predicted
SUMMARY   #length 235 #molecular-weight 25806 #checksum 8212

Query Match   6.5%; Score 150; DB 2; Length 235;
Best Local Similarity 25.6%; Pred. No. 3.00e-07;
Matches   33; Conservative   32; Mismatches   58; Indels   6; Gaps   6;

Db 107 WLSQRANALLANGMDLKDNQLVVPADGLYLIYSQVLFKGQGCPDYVLLTH-TVSRFAISY 165
        | :|: | :::|| ::| | : :||: |||| : : :: :::|:
Qy 186 WYHDRGWAKISN-MTLSNGKLRVNQDGFYLYANICFRHHETSGSVPTDYIQLMYVVKT 244

Db 166 QEKV-SLLSAIKSPCPKDTPEGAELKPWTEPMYLGGVFQLEKGDLLSAEVNLPKYLDITE 224
        : : :|| :|| | : : | | :: :| || :||
Qy 245 SIKIPSSHNLMKGGSTKNWSGNSEFH-FY-SINVGGFFKLRAGEEISIQVSNPSLLD-PD 301

Db 225 SGQVYFGVI 233
        |||
Qy 302 QDATYFGAF 310

RESULT   10
ENTRY   S21738   #type complete
TITLE   CD40 ligand - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE   31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
         10-Sep-1997
ACCESSIONS S21738
REFERENCE S21738
    #authors Armitage, R.J.; Fanslow, W.C.; Strockbine, L.; Sato, T.A.;
              Clifford, K.N.; Macduff, B.M.; Anderson, D.M.; Gimpel,
              S.D.; Davis-Smith, T.; Maliszewski, C.R.; Clark, E.A.;
              Smith, C.A.; Grabstein, K.H.; Cosman, D.; Spriggs, M.K.
    #journal Nature (1992) 357:80-82
    #title   Molecular and biological characterization of a murine ligand
              for CD40.
    #cross-references MUID:92244364
    #accession S21738
    ##molecule_type mRNA
    ##residues 1-260 ##label ARM
    ##cross-references EMBL:X65453; NID:g50351; PID:g50352
KEYWORDS   glycoprotein; transmembrane protein
FEATURE
    23-46   #domain transmembrane #status predicted #label TMM\
    47-260   #domain extracellular #status predicted #label EXT\
    239   #binding_site carbohydrate (Asn) (covalent) #status
              predicted
SUMMARY   #length 260 #molecular-weight 29396 #checksum 9490

Query Match   6.5%; Score 150; DB 2; Length 260;
Best Local Similarity 29.6%; Pred. No. 3.00e-07;
Matches   48; Conservative   49; Mismatches   47; Indels   18; Gaps   16;

Db 113 QRGDEDPQIAAHV-VSEAN-SNAA-SV-LQ-W-AKKGYYTMKSNLVMLENGKQLTVKREG 166
        ||| || :|: : :| ::: | :| :|||| |||| :::|

```
Qy  155 QRGKPEAQPFAHLTINAASIPSGSHKVTLSSWYHDRGW-AKISNMT-LSNGK-LRVNQDG 211
Db  167 LYYVTYQVTFCSNREPSSQRP--FI-VGLWL-KPSI--GSERILLKAANTHSSSQLCE-Q 219
Qy  212 FYYLYANICF-RHHETSGSVPTDYLQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFH 270
Db  220 -QSVHLGGVFELQAGASVFVNVTEASQVIHRVGFSSFGLLKL 260
Qy  271 FYSINVGGFFKLRAGEEISIQVSNPSLLDPDQDATYFGAFKV 312
```

RESULT 11
ENTRY I54490 #type complete
TITLE tumor necrosis factor - white-footed mouse
ORGANISM #formal_name Peromyscus leucopus #common_name white-footed
          mouse
DATE 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change
      23-Feb-1997
ACCESSIONS I54490
REFERENCE I54490
  #authors Crew, M.D.; Filipowsky, M.E.
  #journal Immunogenetics (1992) 35:351-353
  #title Sequence of the tumor necrosis factor/cachectin (TNF) gene
          from Peromyscus leucopus (family Cricetidae).
  #cross-references MUID:92218012
  #accession I54490
  ##status preliminary; translated from GB/EMBL/DDBJ
  ##molecule_type DNA
  ##residues 1-235 ##label RES
  ##cross-references GB:M59233; NID:g202506; PID:g202507
GENETICS
  #gene P1TNF
  #introns 62/3; 81/1; 97/1
CLASSIFICATION #superfamily tumor necrosis factor
SUMMARY #length 235 #molecular-weight 25822 #checksum 8367

Query Match 6.3%; Score 145; DB 2; Length 235;
Best Local Similarity 26.3%; Pred. No. 1.60e-06;
Matches 31; Conservative 29; Mismatches 53; Indels 5; Gaps 5;

```
Db  118 NGMDLKDNQLVIPADGLYLVYSQVLFKGQGCSSYVLLTH-TVSRFAVSYEDKV-NLLSAI 175
Qy  196 SNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYVVKTSIKIPSSHNLM 255
Db  176 KSPCPKETPEGSELKPWYEPIYLGGVFQLEKGDRLSAEVNLPKYLDFAESGQVYFGVI 233
Qy  256 KGGSTKNWSGNSEFH-FY-SINVGGFFKLRAGEEISIQVSNPSLLD-PDQDATYFGAF 310
```

RESULT 12
ENTRY JQ1344 #type complete
TITLE tumor necrosis factor alpha precursor - horse
ALTERNATE_NAMES cachectin; TNF alpha
ORGANISM #formal_name Equus caballus #common_name domestic horse
DATE 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change
      08-Sep-1997
ACCESSIONS JQ1344
REFERENCE JQ1344
  #authors Su, X.; Morris, D.D.; McGraw, R.A.
  #journal Gene (1991) 107:319-321
  #title Cloning and characterization of gene TNF alpha encoding
          equine tumor necrosis factor alpha.
  #cross-references MUID:92084125
  #accession JQ1344
  ##molecule_type DNA
  ##residues 1-234 ##label SUX
  ##cross-references GB:M64087; NID:g164244; PID:g164245
COMMENT This protein is an important proximal mediator of endotoxemia.
GENETICS
  #gene TNF-alpha
  #introns 62/3; 79/1; 95/1
CLASSIFICATION #superfamily tumor necrosis factor
KEYWORDS cytokine; cytotoxin; glycoprotein; lymphokine; macrophage;

membrane protein

FEATURE
  78-234 #product tumor necrosis factor alpha #status predicted
          #label TUM
SUMMARY #length 234 #molecular-weight 25469 #checksum 7690

Query Match 6.2%; Score 143; DB 2; Length 234;
Best Local Similarity 27.7%; Pred. No. 3.10e-06;
Matches 36; Conservative 33; Mismatches 50; Indels 11; Gaps 7;

```
Db  105 WLSGRANALLANGVKLTDNQLVVPLDGLIYSQVLFKGQ--GCPSTHVL-LTHTISRL 160
Qy  186 WYHDRGWAKISN-MTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYVVKT 244
Db  161 AVSYPSKVNLLSAIKSPCHTESPEQAEAKPWYEPIYLGGVFQLEKGDQLSAEINQPNYLD 220
Qy  245 SIKIPSSHNLM--KGGSTKNWSGNSEFH-FY-SINVGGFFKLRAGEEISIQVSNPSLLD 299
Db  221 FAESGQVYFG 230
Qy  300 -PDQDATYFG 308
```

RESULT 13
ENTRY 1CDAB #type complete
TITLE Cd40 ligand complex with the membrane-bound glycoprotein cd40
        (theoretical model), chain B - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
REFERENCE A51090
  #authors Peitsch, M.C.; Jongeneel, C.V.
  #submission submitted to the Brookhaven Protein Data Bank, July 1992
  #cross-references PDB:1CDA
REFERENCE TN06018
  #authors Peitsch, M.C.; Jongeneel, C.V.
  #journal Int. Immunol. (1993) 5:233
  #title A 3-dimensional model for the cd40 ligand reveals a close
          similarity to the tumor necrosis factors.
REFERENCE S21738
  #authors Armitage, R.J.; Fanslow, W.C.; Strockbine, L.; Sato, T.A.;
            Clifford, K.N.; Macduff, B.M.; Anderson, D.M.; Gimpel,
            S.D.; Davis-smith, T.; Maliszewski, C.R.; Clark, E.A.;
            Smith, C.A.; Grabstein, K.H.; Spriggs, M.K.
  #journal Nature (1992) 357:80
  #title Molecular and biological characterization of a murine ligand
          for cd40.
  #cross-references MUID:92244364
REFERENCE A42310
  #authors Eck, M.J.; Ultsch, M.; Rinderknecht, E.; De Vos, A.M.;
            Sprang, S.R.
  #journal J. Biol. Chem. (1992) 267:2119
  #title The structure of human lymphotoxin (tumor necrosis
          factor-beta) at 1.9 angstroms resolution.
REFERENCE TN06021
  #authors Eck, M.J.; Sprang, S.R.
  #journal J. Biol. Chem. (1989) 264:17595
  #title The structure of tumor necrosis factor-alpha at 2.6 angstroms
          resolution. Implications for receptor binding.
COMMENT Resolution: not applicable
COMMENT Determination: theoretical model
FEATURE
  63-103 #disulfide_bonds
SUMMARY #length 146 #molecular-weight 16020 #checksum 7684

Query Match 6.1%; Score 141; DB 5; Length 146;
Best Local Similarity 28.2%; Pred. No. 6.00e-06;
Matches 42; Conservative 45; Mismatches 48; Indels 14; Gaps 12;

```
Db  7 IAAHVVSEANSNAASVLQW-AKKGYYTMKSNLVMLENGKQLTVKREGLYYVTYQVTFCSN 65
Qy  169 INAASIP-SGSHKVTLSSWYHDRGW-AKISNMT-LSNGK-LRVNQDGFYYLYANICF-RH 223
Db  66 REPSSQRP--FI-VGLWL-KPSI--GSERILLKAANTHSSSQLCE-Q-QSVHLGGVFELQ 117
```

```
Qy   224 HETSGSVPTDYLQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLR 283

Db   118 AGASVFVNVTEASQVIHRVGFSSFGLLKL 146
             || :|:::|:|: :| :||:  || :|:
Qy   284 AGEEISIQVSNPSLLDPDQDATYFGAFKV 312


RESULT   14
ENTRY    1CDAC    #type complete
TITLE    Cd40 ligand complex with the membrane-bound glycoprotein CD40
         (theoretical model), chain C - mouse
ORGANISM A51090
         #formal_name Mus musculus #common_name house mouse
REFERENCE
#authors Peitsch, M.C.; Jongeneel, C.V.
#submission submitted to the Brookhaven Protein Data Bank, July 1992
#cross-references PDB:1CDA
REFERENCE TN006022
#authors Peitsch, M.C.; Jongeneel, C.V.
#journal Int. Immunol. (1993) 5:233
#title A 3-dimensional model for the cd40 ligand reveals a close
       similarity to the tumor necrosis factors.
REFERENCE S21738
#authors Armitage, R.J.; Fanslow, W.C.; Strockbine, L.; Sato, T.A.;
         Clifford, K.N.; Macduff, B.M.; Anderson, D.M.; Gimpel,
         S.D.; Davis-smith, T.; Maliszewski, C.R.; Clark, E.A.;
         Smith, C.A.; Grabstein, K.H.; Spriggs, M.K.
#journal Nature (1992) 357:80
#title Molecular and biological characterization of a murine ligand
       for cd40.
#cross-references MUID:92244364
REFERENCE A42310
#authors Eck, M.J.; Ultsch, M.; Rinderknecht, E.; De Vos, A.M.;
         Sprang, S.R.
#journal J. Biol. Chem. (1992) 267:2119
#title The structure of human lymphotoxin (tumor necrosis
       factor-beta) at 1.9 angstroms resolution.
REFERENCE TN006025
#authors Eck, M.J.; Sprang, S.R.
#journal J. Biol. Chem. (1989) 264:17595
#title The structure of tumor necrosis factor-alpha at 2.6 angstroms
       resolution. Implications for receptor binding.
COMMENT  Resolution: not applicable
COMMENT  Determination: theoretical model
FEATURE
  63-103           #disulfide_bonds
SUMMARY  #length 146 #molecular-weight 16020 #checksum 7684
Query Match     6.1%;    Score 141; DB 5; Length 146;
Best Local Similarity 28.2%;    Pred. No. 6.00e-06;
Matches  42; Conservative  45; Mismatches 48; Indels 14; Gaps 12;

Db     7 IAAHVVSEANSNAASVLQW-AKKGYYTMKSNLVMLENGKQLTVKREGLYYYTQVTFCSN 65
           |:|::|:::||: .:| : :|||:|||:::: |:::::|:::|::| :|:-  :
Qy   169 INAASIP-SGSHKVTLSSWYHDRGW-AKISNMT-LSNGK-LRVNQDGFYLYANICF-RH 223

Db    66 REPSSQRP--FI-VGLWL-KPSI--GSERILLKAANTHSSSQLCE-Q-QSVHLGGVFELQ 117
           :|:|: :    ::||:: ||: ! | :::|| : :: |:|| |:!  !:::| :! ::
Qy   224 HETSGSVPTDYLQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLR 283

Db   118 AGASVFVNVTEASQVIHRVGFSSFGLLKL 146
             || :|:::|:|: :| :||:  || :|:
Qy   284 AGEEISIQVSNPSLLDPDQDATYFGAFKV 312


RESULT   15
ENTRY    1CDAA    #type complete
TITLE    Cd40 ligand complex with the membrane-bound glycoprotein Cd40
         (theoretical model), chain A - mouse
ORGANISM A51090
         #formal_name Mus musculus #common_name house mouse
REFERENCE
#authors Peitsch, M.C.; Jongeneel, C.V.
#submission submitted to the Brookhaven Protein Data Bank, July 1992
#cross-references PDB:1CDA
REFERENCE TN006014
#authors Peitsch, M.C.; Jongeneel, C.V.
#journal Int. Immunol. (1993) 5:233
#title A 3-dimensional model for the cd40 ligand reveals a close
       similarity to the tumor necrosis factors.
REFERENCE S21738
#authors Armitage, R.J.; Fanslow, W.C.; Strockbine, L.; Sato, T.A.;
         Clifford, K.N.; Macduff, B.M.; Anderson, D.M.; Gimpel,
         S.D.; Davis-smith, T.; Maliszewski, C.R.; Clark, E.A.;
         Smith, C.A.; Grabstein, K.H.; Spriggs, M.K.
#journal Nature (1992) 357:80
#title Molecular and biological characterization of a murine ligand
       for cd40.
#cross-references MUID:92244364
REFERENCE A42310
#authors Eck, M.J.; Ultsch, M.; Rinderknecht, E.; De Vos, A.M.;
         Sprang, S.R.
#journal J. Biol. Chem. (1992) 267:2119
#title The structure of human lymphotoxin (tumor necrosis
       factor-beta) at 1.9 angstroms resolution.
REFERENCE TN006017
#authors Eck, M.J.; Sprang, S.R.
#journal J. Biol. Chem. (1989) 264:17595
#title The structure of tumor necrosis factor-alpha at 2.6 angstroms
       resolution. Implications for receptor binding.
COMMENT  Resolution: not applicable
COMMENT  Determination: theoretical model
FEATURE
  33-35,7-13,139-146,          #region beta sheet\
  52-65,104-116
  39-41,45-48,                 #region beta sheet\
  119-126,74-82,
  85-95                        #disulfide_bonds
  63-103
SUMMARY  #length 146 #molecular-weight 16020 #checksum 7684
Query Match     6.1%;    Score 141; DB 5; Length 146;
Best Local Similarity 28.2%;    Pred. No. 6.00e-06;
Matches  42; Conservative  45; Mismatches 48; Indels 14; Gaps 12;

Db     7 IAAHVVSEANSNAASVLQW-AKKGYYTMKSNLVMLENGKQLTVKREGLYYYTQVTFCSN 65
           |:|::|:::||: .:| : :|||:|||:::: |:::::|:::|::| :|:-  :
Qy   169 INAASIP-SGSHKVTLSSWYHDRGW-AKISNMT-LSNGK-LRVNQDGFYLYANICF-RH 223

Db    66 REPSSQRP--FI-VGLWL-KPSI-GSERILLKAANTHSSSQLCE-Q-QSVHLGGVFELQ 117
           :|:|: :    ::||:: ||: ! | :::|| : :: |:|| |:!  !:::| :! ::
Qy   224 HETSGSVPTDYLQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLR 283

Db   118 AGASVFVNVTEASQVIHRVGFSSFGLLKL 146
             || :|:::|:|: :| :||:  || :|:
Qy   284 AGEEISIQVSNPSLLDPDQDATYFGAFKV 312


Search completed: Fri Nov 13 17:15:14 1998
Job time : 87 secs.
```

```
 __    __  ____   _____  ____    ____  _   _
|  \  /  ||  _ \ / ____||  _ \  / ___|| | | |
|   \/   || |_) | (___  | |_) || |    | |_| |
| |\  /| ||  __/ \___ \ |  _ < | |    |  _  |  (TM)
| | \/ | || |    ____) || | \ \| |___ | | | |
|_|    |_||_|   |_____/ |_|  \_\\____||_| |_|
```

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

on on: Fri Nov 13 17:19:15 1998;   MasPar time 3.46 Seconds
                                           647.079 Million cell updates/sec
Tabular output not generated.

Title: >US-08-989-362-2
Description: (1-316) from US08989362.pep
Perfect Score: 2294
Sequence: 1 MRRASRDYGKYLRSSEMGS.........LLDPDQDATYFGAFKVODID 316

Scoring table: PAM 150
                Gap 11

Searched: 77309 seqs, 7078906 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database: a-issued
                1:5_COMB 2:PCT9_COMB 3:backfiles1

Statistics: Mean 31.764; Variance 144.069; scale 0.220

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

| Result No. | Score | % Query Match | Query Length | DB | ID | Description | Pred. No. |
|---|---|---|---|---|---|---|---|
| 1 | 326 | 14.2 | 281 | 1 | US-08-670- | Sequence 2, Applicatio | 9.86e-19 |
| 2 | 326 | 14.2 | 281 | 1 | PCT-US96-1 | Sequence 2, Applicatio | 9.86e-19 |
| 3 | 282 | 12.3 | 291 | 1 | US-08-670- | Sequence 6, Applicatio | 6.84e-15 |
| 4 | 282 | 12.3 | 291 | 2 | PCT-US96-1 | Sequence 2, Applicatio | 6.84e-15 |
| 5 | 175 | 7.6 | 281 | 2 | PCT-US95-0 | Sequence 2, Applicatio | 7.32e-06 |
| 6 | 170 | 7.4 | 279 | 2 | PCT-US95-0 | Sequence 5, Applicatio | 1.86e-06 |
| 7 | 167 | 7.3 | 151 | 1 | US-07-940- | Sequence 3, Applicatio | 3.25e-05 |
| 8 | 165 | 7.2 | 261 | 1 | US-08-360- | Sequence 2, Applicatio | 4.71e-05 |
| 9 | 165 | 7.2 | 261 | 1 | US-08-446- | Sequence 4, Applicatio | 4.71e-05 |
| 10 | 165 | 7.2 | 261 | 1 | US-07-940- | Sequence 2, Applicatio | 4.71e-05 |
| 11 | 165 | 7.2 | 261 | 1 | US-08-184- | Sequence 8, Applicatio | 4.71e-05 |
| 12 | 165 | 7.2 | 261 | 2 | PCT-US93-1 | Sequence 4, Applicatio | 4.71e-05 |
| 13 | 165 | 7.2 | 273 | 1 | US-08-446- | Sequence 11, Applicati | 4.71e-05 |
| 14 | 162 | 7.1 | 157 | 2 | PCT-US93-0 | Sequence 9, Applicatio | 8.20e-05 |
| 15 | 154 | 6.7 | 151 | 1 | US-07-940- | Sequence 4, Applicatio | 3.57e-04 |
| 16 | 148 | 6.5 | 157 | 2 | PCT-US93-0 | Sequence 7, Applicatio | 1.06e-03 |
| 17 | 148 | 6.5 | 157 | 2 | PCT-US93-0 | Sequence 8, Applicatio | 1.06e-03 |
| 18 | 142 | 6.2 | 260 | 1 | US-08-446- | Sequence 6, Applicatio | 3.14e-03 |
| 19 | 142 | 6.2 | 260 | 1 | PCT-US93-1 | Sequence 6, Applicatio | 3.14e-03 |
| 20 | 141 | 6.1 | 158 | 1 | US-07-794- | Sequence 6, Applicatio | 3.76e-03 |
| 21 | 141 | 6.1 | 158 | 1 | US-08-397- | Sequence 6, Applicatio | 3.76e-03 |
| 22 | 136 | 5.9 | 150 | 1 | US-07-668- | Sequence 12, Applicati | 9.19e-03 |
| 23 | 136 | 5.9 | 151 | 1 | US-07-668- | Sequence 26, Applicati | 9.19e-03 |
| 24 | 136 | 5.9 | 155 | 1 | US-07-994- | Sequence 94, Applicati | 9.19e-03 |
| 25 | 135 | 5.9 | 157 | 2 | PCT-US93-0 | Sequence 4, Applicatio | 1.10e-02 |
| 26 | 136 | 5.9 | 158 | 1 | US-07-994- | Sequence 99, Applicatio | 9.19e-03 |
| 27 | 135 | 5.9 | 158 | 1 | US-08-397- | Sequence 7, Applicatio | 1.10e-02 |
| 28 | 135 | 5.9 | 158 | 1 | US-07-794- | Sequence 76, Applicatio | 1.10e-02 |
| 29 | 134 | 5.8 | 155 | 1 | US-07-994- | Sequence 3, Applicatio | 1.31e-02 |
| 30 | 133 | 5.8 | 155 | 1 | US-07-994- | Sequence 82, Applicatio | 1.57e-02 |
| 31 | 134 | 5.8 | 158 | 1 | US-07-994- | Sequence 8, Applicatio | 1.31e-02 |
| 32 | 133 | 5.8 | 158 | 1 | US-08-397- | Sequence 8, Applicatio | 1.57e-02 |
| 33 | 133 | 5.7 | 130 | 1 | US-07-794- | Sequence 75, Applicati | 1.57e-02 |
| 34 | 130 | 5.7 | 154 | 1 | US-07-994- | Sequence 92, Applicati | 2.66e-02 |
| 35 | 131 | 5.7 | 155 | 1 | US-07-994- | Sequence 13, Applicatio | 2.23e-02 |
| 36 | 130 | 5.7 | 155 | 1 | US-07-994- | Sequence 5, Applicatio | 2.66e-02 |
| 37 | 130 | 5.7 | 157 | 1 | US-07-994- | Sequence 81, Applicatio | 2.66e-02 |
| 38 | 130 | 5.7 | 157 | 1 | US-07-940- | Sequence 5, Applicatio | 2.66e-02 |
| 39 | 130 | 5.7 | 158 | 1 | US-07-994- | Sequence 97, Applicatio | 2.66e-02 |
| 40 | 131 | 5.7 | 158 | 1 | US-07-994- | Sequence 10, Applicatio | 2.23e-02 |
| 41 | 129 | 5.6 | 139 | 1 | US-07-994- | Sequence 9, Applicatio | 3.17e-02 |
| 42 | 129 | 5.6 | 145 | 1 | US-07-994- | Sequence 93, Applicatio | 3.17e-02 |
| 43 | 129 | 5.6 | 155 | 1 | US-07-994- | Sequence 79, Applicati | 3.17e-02 |
| 44 | 129 | 5.6 | 155 | 1 | US-07-994- | Sequence 98, Applicati | 3.17e-02 |
| 45 | 129 | 5.6 | 158 | 1 | US-07-994- | Sequence 98, Applicati | 3.17e-02 |

ALIGNMENTS

```
RESULT 1
ID   US-08-670-354-2          STANDARD;       PRT;      281 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 2, Application US/08670354
XX
CC   Sequence 2, Application US/08670354
CC   Patent No. 5763223
CC   GENERAL INFORMATION:
CC     APPLICANT:  Steven R. Wiley and
CC     APPLICANT:  Raymond G. Goodwin.
CC     TITLE OF INVENTION:  Cytokine That Induces Apoptosis
CC     NUMBER OF SEQUENCES:  9
CC     CORRESPONDENCE ADDRESS:
CC       ADDRESSEE:  Kathryn A. Anderson, Immunex Corporation
CC       STREET:  51 University Street
CC       CITY:  Seattle
CC       STATE:  WA
CC       COUNTRY:  USA
CC       ZIP:  98101
CC     COMPUTER READABLE FORM:
CC       MEDIUM TYPE:  Floppy disk
CC       COMPUTER:  Apple Macintosh
CC       OPERATING SYSTEM:  Apple 7.5.2
CC       SOFTWARE:  Microsoft Word, Version 6.0.1
CC     CURRENT APPLICATION DATA:
CC       APPLICATION NUMBER:  US/08/670,354
CC       FILING DATE:  25-JUN-1996
CC       CLASSIFICATION:  435
CC     PRIOR APPLICATION DATA:
CC       APPLICATION NUMBER:  US 08/496,632
CC       FILING DATE:  29-JUN-1995
CC       CLASSIFICATION:  435
CC     PRIOR APPLICATION DATA:
CC       APPLICATION NUMBER:  US 08/548,368
CC       FILING DATE:  01-NOV-1995
CC       CLASSIFICATION:  435
CC     ATTORNEY/AGENT INFORMATION:
CC       NAME:  Anderson, Kathryn A.
CC       REGISTRATION NUMBER:  32,172
CC       REFERENCE/DOCKET NUMBER:  2835-B
CC     TELECOMMUNICATION INFORMATION:
CC       TELEPHONE:  (206) 587-0430
CC       TELEFAX:  (206) 233-0644
```

```
CC     TELEX: 756822
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC   LENGTH: 281 amino acids
CC   TYPE: amino acid
CC   TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE  281 AA; 32509 MW; 420741 CN;

Query Match      14.2%;   Score 326;  DB 1;  Length 281;
Best Local Similarity   35.9%;   Pred. No. 9.86e-19;
Matches   47;  Conservative   33;  Mismatches  49;  Indels  2;  Gaps  2;

Db  151 INSWESSRSGHSFLSNLHLRNGELVIHEKGFYYIYSQTYFRFQEEIKENTKNDKQMVQYI 210
            ::||| -|-::||-:||:||:|:|| |||||:|: || :|: |:
Qy  183 LSSWYHDR-GWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYV 241
          -|-:|| ::|||: |: |||: ::|-|||:||| :|:|:|| ||:|
Db  211 YKYT-SYPDPILLMKSARNSCWSKDAEYGLYSIYQGGIFELKENDRIFVSVTNEHLIDMD 269
          - |-:||| |: ||::: |||: :|||||:|||:|||| :-|: :||| |:|
Qy  242 VKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDPD 301

Db  270 HEASFFGAFLV 280
          ::|:||||
Qy  302 QDATYFGAFKV 312

RESULT  2
ID  PCT-US96-10895-2    STANDARD;    PRT;    281 AA.
XX
AC  xxxxxx
XX
DT
XX
DE  Sequence 2, Application PC/TUS9610895
XX
CC Sequence 2, Application PC/TUS9610895
CC GENERAL INFORMATION:
CC   APPLICANT: Immunex Corporation.
CC   TITLE OF INVENTION: Cytokine That Induces Apoptosis
CC   NUMBER OF SEQUENCES: 9
CC   CORRESPONDENCE ADDRESS:
CC     ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
CC     STREET: 51 University Street
CC     CITY: Seattle
CC     STATE: WA
CC     COUNTRY: USA
CC     ZIP: 98101
CC   COMPUTER READABLE FORM:
CC     MEDIUM TYPE: Floppy disk
CC     COMPUTER: Apple Macintosh
CC     OPERATING SYSTEM: Apple 7.5.2
CC     SOFTWARE: Microsoft Word, Version 6.0.1
CC   CURRENT APPLICATION DATA:
CC     APPLICATION NUMBER: PCT/US96/10895
CC     FILING DATE: 25-JUN-1996
CC     CLASSIFICATION:
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER: US 08/496,632
CC     FILING DATE: 29-JUN-1995
CC     CLASSIFICATION:
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER: US 08/548,368
CC     FILING DATE: 01-NOV-1995
CC     CLASSIFICATION:
CC   ATTORNEY/AGENT INFORMATION:
CC     NAME: Anderson, Kathryn A.
CC     REGISTRATION NUMBER: 32,172
CC     REFERENCE/DOCKET NUMBER: 2835-WO
CC   TELECOMMUNICATION INFORMATION:
CC     TELEPHONE: (206) 587-0430
CC     TELEFAX: (206) 233-0644
CC     TELEX: 756822
CC INFORMATION FOR SEQ ID NO: 2:
```

```
CC SEQUENCE CHARACTERISTICS:
CC   LENGTH: 281 amino acids
CC   TYPE: amino acid
CC   TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE  281 AA; 32509 MW; 420741 CN;

Query Match      14.2%;   Score 326;  DB 2;  Length 281;
Best Local Similarity   35.9%;   Pred. No. 9.86e-19;
Matches   47;  Conservative   33;  Mismatches  49;  Indels  2;  Gaps  2;

Db  151 INSWESSRSGHSFLSNLHLRNGELVIHEKGFYYIYSQTYFRFQEEIKENTKNDKQMVQYI 210
            ::||| -|-::||-:||:||:|:|| |||||:|: || :|: |:
Qy  183 LSSWYHDR-GWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYV 241
          -|-:|| ::|||: |: |||: ::|-|||:||| :|:|:|| ||:|
Db  211 YKYT-SYPDPILLMKSARNSCWSKDAEYGLYSIYQGGIFELKENDRIFVSVTNEHLIDMD 269
          - |-:||| |: ||::: |||: :|||||:|||:|||| :-|: :||| |:|
Qy  242 VKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDPD 301

Db  270 HEASFFGAFLV 280
          ::|:||||
Qy  302 QDATYFGAFKV 312

RESULT  3
ID  US-08-670-354-6    STANDARD;    PRT;    291 AA.
XX
AC  xxxxxx
XX
DT
XX
DE
XX
CC Sequence 6, Application US/08670354
CC Patent No.5763223
CC GENERAL INFORMATION:
CC   APPLICANT: Steven R. Wiley and
CC   APPLICANT: Raymond G. Goodwin.
CC   TITLE OF INVENTION: Cytokine That Induces Apoptosis
CC   NUMBER OF SEQUENCES: 9
CC   CORRESPONDENCE ADDRESS:
CC     ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
CC     STREET: 51 University Street
CC     CITY: Seattle
CC     STATE: WA
CC     COUNTRY: USA
CC     ZIP: 98101
CC   COMPUTER READABLE FORM:
CC     MEDIUM TYPE: Floppy disk
CC     COMPUTER: Apple Macintosh
CC     OPERATING SYSTEM: Apple 7.5.2
CC     SOFTWARE: Microsoft Word, Version 6.0.1
CC   CURRENT APPLICATION DATA:
CC     APPLICATION NUMBER: US/08/670,354
CC     FILING DATE: 25-JUN-1996
CC     CLASSIFICATION: 435
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER: US 08/496,632
CC     FILING DATE: 29-JUN-1995
CC     CLASSIFICATION: 435
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER: US 08/548,368
CC     FILING DATE: 01-NOV-1995
CC     CLASSIFICATION: 435
CC   ATTORNEY/AGENT INFORMATION:
CC     NAME: Anderson, Kathryn A.
CC     REGISTRATION NUMBER: 32,172
CC     REFERENCE/DOCKET NUMBER: 2835-B
CC   TELECOMMUNICATION INFORMATION:
CC     TELEPHONE: (206) 587-0430
CC     TELEFAX: (206) 233-0644
CC     TELEX: 756822
CC INFORMATION FOR SEQ ID NO: 6:
```

```
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH: 291 amino acids
CC          TYPE: amino acid
CC          TOPOLOGY: linear
CC        MOLECULE TYPE: protein
SQ        SEQUENCE  291 AA; 33477 MW; 453904 CN;

Query Match              12.3%;  Score 282; DB 1; Length 291;
Best Local Similarity    35.6%;  Pred. No. 6.84e-15;
Matches  48;  Conservative  37;  Mismatches 42;  Indels  8;  Gaps  5;

Db  155  IESWESSRKGHSFLNHVLFRNGELVIEQEGLYYIYSQTYFRFQEAEDASKMVSKDKVRTK 214
             |  ||  |  :::  ||  |||||||:  ||  :||  :|  :| :|:
Qy  183  LSSWYHDR-GWAKISNMTLSNGKLRVNQDGFYYLYANICFRHET-SGS--VPTDYL--- 235
             |:  |  :|||  ||  ::|:  ||  |  ::||:  |||  ||||:  |  :  ||
Db  215  QLVQYIYKYT-SYPDPIVLMKSARNSCWSRDAEYGLYSIYQGGLFELKKNDRIFVSVTNE 273
             |:  :|::  :  |||:  |  ||  ||:|||  |||  |  |  :  ||
Qy  236  QLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNP 295

Db  274  HLMDLDQEASFFGAF 288
              :|  ||  ||:::|||
Qy  296  SLLDPDQDATYFGAF 310


RESULT   4
ID   PCT-US96-10895-6      STANDARD;    PRT;    291 AA.
XX
AC   xxxxxx
XX
XX
DT
XX
DE   Sequence 6, Application PC/TUS9610895
XX
CC   Sequence 6, Application PC/TUS9610895
CC   GENERAL INFORMATION:
CC     APPLICANT: Immunex Corporation.
CC     TITLE OF INVENTION: Cytokine That Induces Apoptosis
CC     NUMBER OF SEQUENCES: 9
CC     CORRESPONDENCE ADDRESS:
CC       ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
CC       STREET: 51 University Street
CC       CITY: Seattle
CC       STATE: WA
CC       COUNTRY: USA
CC       ZIP: 98101
CC     COMPUTER READABLE FORM:
CC       MEDIUM TYPE: Floppy disk
CC       COMPUTER: Apple Macintosh
CC       OPERATING SYSTEM: Apple 7.5.2
CC       SOFTWARE: Microsoft Word, Version 6.0.1
CC     CURRENT APPLICATION DATA:
CC       APPLICATION NUMBER: PCT/US96/10895
CC       FILING DATE: 25-JUN-1996
CC       CLASSIFICATION:
CC     PRIOR APPLICATION DATA:
CC       APPLICATION NUMBER: US 08/496,632
CC       FILING DATE: 29-JUN-1995
CC       CLASSIFICATION:
CC     PRIOR APPLICATION DATA:
CC       APPLICATION NUMBER: US 08/548,368
CC       FILING DATE: 01-NOV-1995
CC       CLASSIFICATION:
CC     ATTORNEY/AGENT INFORMATION:
CC       NAME: Anderson, Kathryn A.
CC       REGISTRATION NUMBER: 32,172
CC       REFERENCE/DOCKET NUMBER: 2835-WO
CC     TELECOMMUNICATION INFORMATION:
CC       TELEPHONE:  (206) 587-0430
CC       TELEFAX: (206) 233-0644
CC       TELEX: 756822
CC   INFORMATION FOR SEQ ID NO: 6:
CC     SEQUENCE CHARACTERISTICS:
CC       LENGTH: 291 amino acids
CC       TYPE: amino acid
CC       TOPOLOGY: linear
CC       MOLECULE TYPE: protein
SQ       SEQUENCE  291 AA; 33477 MW; 453904 CN;

Query Match              12.3%;  Score 282; DB 2; Length 291;
Best Local Similarity    35.6%;  Pred. No. 6.84e-15;
Matches  48;  Conservative  37;  Mismatches 42;  Indels  8;  Gaps  5;

Db  155  IESWESSRKGHSFLNHVLFRNGELVIEQEGLYYIYSQTYFRFQEAEDASKMVSKDKVRTK 214
             |  ||  |  :::  ||  |||||||:  ||  :||  :|  :| :|:
Qy  183  LSSWYHDR-GWAKISNMTLSNGKLRVNQDGFYYLYANICFRHET-SGS--VPTDYL--- 235
             |:  |  :|||  ||  ::|:  ||  |  ::||:  |||  ||||:  |  :  ||
Db  215  QLVQYIYKYT-SYPDPIVLMKSARNSCWSRDAEYGLYSIYQGGLFELKKNDRIFVSVTNE 273
             |:  :|::  :  |||:  |  ||  ||:|||  |||  |  |  :  ||
Qy  236  QLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNP 295

Db  274  HLMDLDQEASFFGAF 288
              :|  ||  ||:::|||
Qy  296  SLLDPDQDATYFGAF 310


RESULT   5
ID   PCT-US95-00362-2      STANDARD;    PRT;    281 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 2, Application PC/TUS9500362
XX
CC   Sequence 2, Application PC/TUS9500362
CC   GENERAL INFORMATION:
CC     APPLICANT: IMMUNEX CORPORATION
CC     TITLE OF INVENTION: Ligand That Binds Fas Antigen
CC     NUMBER OF SEQUENCES: 5
CC     CORRESPONDENCE ADDRESS:
CC       ADDRESSEE: Immunex Corporation
CC       STREET: 51 University Street
CC       CITY: Seattle
CC       STATE: WA
CC       COUNTRY: US
CC       ZIP: 98101
CC     COMPUTER READABLE FORM:
CC       MEDIUM TYPE: Floppy disk
CC       COMPUTER: Apple Macintosh
CC       OPERATING SYSTEM: Apple 7.1
CC       SOFTWARE: Microsoft Word, Version 5.1a
CC     CURRENT APPLICATION DATA:
CC       APPLICATION NUMBER: PCT/US95/00362
CC       FILING DATE: 06-JAN-1995
CC       CLASSIFICATION:
CC     PRIOR APPLICATION DATA:
CC       APPLICATION NUMBER: US 08/179,138
CC       FILING DATE: 07-JAN-1994
CC     PRIOR APPLICATION DATA:
CC       APPLICATION NUMBER: US 08/190,559
CC       FILING DATE: 01-FEB-1994
CC     ATTORNEY/AGENT INFORMATION:
CC       NAME: Anderson, Kathryn A.
CC       REGISTRATION NUMBER: 32,172
CC       REFERENCE/DOCKET NUMBER: 2805-WO
CC   INFORMATION FOR SEQ ID NO: 2:
CC     SEQUENCE CHARACTERISTICS:
CC       LENGTH: 281 amino acids
CC       TYPE: amino acid
CC     MOLECULE TYPE: protein
SQ     SEQUENCE  281 AA; 31485 MW; 464188 CN;

Query Match              7.6%;  Score 175; DB 2; Length 281;
Best Local Similarity    23.9%;  Pred. No. 7.32e-06;
Matches  44;  Conservative  55;  Mismatches 75;  Indels 10;  Gaps  8;
```

```
Db 107 LQKELAELR-ESTSQMHTASSLE-KQIGHPSPPEKKELRKVAHLTGKSNSRSMPLEWED 164
       ||  ||.  .::.:|.  .:|:.: |  |  :|  | ...:||.. .|:.:|.:.: |
Qy 130 LQHIVGPQRFSGAPAMMEGSWLDVAQRGKPEAQPFAHLTINAASIPSGSHKVTLS-SWYH 188

Db 165 TYGIVLLSGVKYKKGGLVINETGLYFVYSKVYFRGQ-SCNNLP--LSHKVYM-RNS-KY 218
       ..:.:.| .:||||:||||||||:||:|||||||| |:||   .:.| :| .|: |
Qy 189 DRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHETSGSVPTDYLQLMVYVVKTSIKI 248

Db 219 PQDL-VMEGKMMSYCTTGQMWARSSYLGAVFNLTSADHLYVNVSELSLVNFEESQTFFG 277
       |:|| .:  ::|.|| |.  ||:||||||||||||||||  |: | :||| |||:||
Qy 249 PSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDPDQDATYFG 308

Db 278 LYKL 281
       :|:
Qy 309 AFKV 312


RESULT 6
ID   PCT-US95-00362-5      STANDARD;     PRT;      279 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 5, Application PC/TUS9500362
XX
CC   GENERAL INFORMATION:
CC    APPLICANT: IMMUNEX CORPORATION
CC    TITLE OF INVENTION: Ligand That Binds Fas Antigen
CC    NUMBER OF SEQUENCES: 5
CC    CORRESPONDENCE ADDRESS:
CC     ADDRESSEE: Immunex Corporation
CC     STREET: 51 University Street
CC     CITY: Seattle
CC     STATE: WA
CC     COUNTRY: US
CC     ZIP: 98101
CC   COMPUTER READABLE FORM:
CC     MEDIUM TYPE: Floppy disk
CC     COMPUTER: Apple Macintosh
CC     OPERATING SYSTEM: Apple 7.1
CC     SOFTWARE: Microsoft Word, Version 5.1a
CC   CURRENT APPLICATION DATA:
CC     APPLICATION NUMBER: PCT/US95/00362
CC     FILING DATE: 06-JAN-1995
CC     CLASSIFICATION:
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER: US 08/179,138
CC     FILING DATE: 07-JAN-1994
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER: US 08/190,559
CC     FILING DATE: 01-FEB-1994
CC   ATTORNEY/AGENT INFORMATION:
CC     NAME: Anderson, Kathryn A.
CC     REGISTRATION NUMBER: 32,172
CC     REFERENCE/DOCKET NUMBER: 2805-WO
CC   INFORMATION FOR SEQ ID NO: 5:
CC   SEQUENCE CHARACTERISTICS:
CC     LENGTH: 279 amino acids
CC     TYPE: amino acid
CC     TOPOLOGY: linear
CC   MOLECULE TYPE: protein
SQ   SEQUENCE 279 AA; 31462 MW; 446260 CN;

Query Match      7.4%;    Score 170;   DB 2;   Length 279;
Best Local Similarity 25.2%;   Pred. No. 1.86e-05;
Matches   34;   Conservative  41;   Mismatches 52;   Indels  8;   Gaps  7;

Db 153 SRSIPLE-WEDTYGTALISGVKYKKGGLVINETGLYFVYSKVYFRGQ-SCNNQPLN--H- 207
Qy 178 SHKVTLSSWHDRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQL 237
```

```
Db 208 KVYM-RNS-KYPEDLVLMEEKRLNYFT-TGQIWAHSSYLGAVFNLTSADHLYVNISQLSL 264
       :||:.||  .:.|:. :||  | |  |   :| |:|| .:|||||||||.||:|:|.| ||
Qy 238 MVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSL 297

Db 265 INFEESKTFFGLYKL 279
       ::.:| |:||.:|:
Qy 298 LDPDQDATYFGAFKV 312


RESULT 7
ID   US-07-940-605A-3      STANDARD;     PRT;      151 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 3, Application US/07940605A
XX
CC   Sequence 3, Application US/07940605A
CC   Patent No. 5540926
CC   GENERAL INFORMATION:
CC    APPLICANT: ARUFFO, ALEJANDRO
CC    APPLICANT: HOLLENBAUGH, DIANE
CC    APPLICANT: LEDBETTER, JEFFREY A.
CC    TITLE OF INVENTION: SOLUBLE LIGANDS FOR CD40
CC    NUMBER OF SEQUENCES: 15
CC    CORRESPONDENCE ADDRESS:
CC     ADDRESSEE: Pennie & Edmonds
CC     STREET: 1155 Avenue of the Americas
CC     CITY: New York
CC     STATE: New York
CC     COUNTRY: U.S.A.
CC     ZIP: 10036-2711
CC   COMPUTER READABLE FORM:
CC     MEDIUM TYPE: Floppy disk
CC     COMPUTER: IBM PC Compatible
CC     OPERATING SYSTEM: PC-DOS/MS-DOS
CC     SOFTWARE: PatentIn Release #1.0, Version #1.25
CC   CURRENT APPLICATION DATA:
CC     APPLICATION NUMBER: US/07/940,605A
CC     FILING DATE: 04-SEP-1992
CC     CLASSIFICATION: 530
CC   ATTORNEY/AGENT INFORMATION:
CC     NAME: Misrock, S. Leslie
CC     REGISTRATION NUMBER: 18,872
CC     REFERENCE/DOCKET NUMBER: 5524-184
CC   TELECOMMUNICATION INFORMATION:
CC     TELEPHONE: 212 790-9090
CC     TELEFAX: 212 869-8864/9741
CC     TELEX: 66141 PENNIE
CC   INFORMATION FOR SEQ ID NO:    3:
CC   SEQUENCE CHARACTERISTICS:
CC     LENGTH: 151 amino acids
CC     TYPE: amino acid
CC     TOPOLOGY: linear
CC   MOLECULE TYPE: protein
SQ   SEQUENCE 151 AA; 16431 MW; 124917 CN;

Query Match      7.3%;    Score 167;   DB 1;   Length 151;
Best Local Similarity 29.5%;   Pred. No. 3.25e-05;
Matches   44;   Conservative  46;   Mismatches 45;   Indels 14;   Gaps  12;

Db 12  IAAHVISEASSKTTSVLQW-AEKGYYTMSNNLVTLENGKQLTVKRQGLYYIYAQVTFCSN 70
       | |  ||| :| :  :| || : ||| |||:| ..::||.|| :. |.:|.::||.||.||
Qy 169 INAASIPSGSHKVT-LSSWYHDRGWAKISN-M-TLSNGK-LRVNQDGFYYLYANICF-RH 223

Db 71  REASSQAP--FIASL-CL-KSPGRF--ERILLRAANTHS-SAKLGGQ-QSIHLGGVFELQ 122
       ::|::  | .:||   || :| .  : |||| .::. :|.. | |.: : ||| |||:||
Qy 224 HETSGSVPTDYLQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLR 283

Db 123 PGASVFVNVTDPSQVSHGTGFTSFGLLKL 151
       .|:  | ||  |:| :    || || .|:
```

Qy   284 AGEEISIQVSNPSLLDPDQDATYFGAFKV 312

RESULT 8
ID   US-08-360-923A-2   STANDARD;   PRT;   261 AA.
XX
AC   xxxxxx
XX
DT
XX
DE
XX
CC   Sequence 2, Application US/08360923A
CC   Patent No. 5574492
CC   GENERAL INFORMATION:
CC   APPLICANT: ARMITAGE, RICHARD
CC   APPLICANT: FANSLOW, WILLIAM
CC   APPLICANT: LONGO, DAN L.
CC   APPLICANT: MURPHY, WILLIAM
CC   TITLE OF INVENTION: METHOD OF PREVENTING OR TREATING
CC   TITLE OF INVENTION: DISEASE CHARACTERIZED BY NEOPLASTIC CELLS
CC   TITLE OF INVENTION: EXPRESSING CD40
CC   NUMBER OF SEQUENCES: 4
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE: IMMUNEX CORPORATION
CC   STREET: 51 UNIVERSITY STREET
CC   CITY: SEATTLE
CC   STATE: WASHINGTON
CC   COUNTRY: USA
CC   ZIP: 98101
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE: Floppy disk
CC   COMPUTER: Macintosh
CC   OPERATING SYSTEM: Apple Macintosh System 7.1
CC   SOFTWARE: Microsoft Word for Macintosh, Version #5.1a
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: US/08/360,923A
CC   FILING DATE: December 21, 1994
CC   CLASSIFICATION: 424
CC   PRIOR APPLICATION DATA:
CC   APPLICATION NUMBER: USSN 08/172,664
CC   FILING DATE: December 23, 1993
CC   CLASSIFICATION: 424
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME: Perkins, Patricia A.
CC   REGISTRATION NUMBER: 34,693
CC   REFERENCE/DOCKET NUMBER: 2818-A
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE: (206)587-0430
CC   TELEFAX: (206)233-0644
CC   INFORMATION FOR SEQ ID NO: 2:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH: 261 amino acids
CC   TYPE: amino acid
CC   TOPOLOGY: linear
CC   MOLECULE TYPE: protein
SQ   SEQUENCE 261 AA; 29273 MW; 363115 CN;

Query Match
Best Local Similarity   7.2%;   Score 165; DB 1; Length 261;
Matches   44; Conservative   47; Mismatches   44; Indels   14; Gaps   12;

Db   122 IAAHVISEASSKTTSVLQW-AEKGYYTMSNNLVTLENGKQLTVKRQGLYIYAQVTFCSN 180
          | |::||:.:| |:| ::| ||:|| ||| ||||:::|::||:| :|:::||:|| :
Qy   169 INAASIPSGSHKVT-LSSWYHDRGWAKISN-M-TLSNGK-LRVNQDGFYLYANICF-RH 223

Db   181 REASSQAP--FIASL-CL-KSPGRF--ERILLRAANTHS-SAKPCGQ-QSIHLGGVFELQ 232
          :|:|| | : | :: | :|:::| | | ::|| ||| :
Qy   224 HETSGSVPTDYIQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLR 283

Db   233 PGASVFVNTDPSQVSHGTGFTSFGLLKL 261
          :| ::|:||:.:: | ||::.:|:
Qy   284 AGEEISIQVSNPSLLDPDQDATYFGAFKV 312

RESULT 9
ID   US-08-446-922-4   STANDARD;   PRT;   261 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 4, Application US/08446922
XX
CC   Sequence 4, Application US/08446922
CC   Patent No. 5716805
CC   GENERAL INFORMATION:
CC   APPLICANT: Spriggs, Melanie
CC   APPLICANT: Srinivasan, Subhashini
CC   TITLE OF INVENTION: Methods of Preparing Soluble, Oligomeric
CC   TITLE OF INVENTION: Proteins
CC   NUMBER OF SEQUENCES: 11
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE: Immunex Corporation
CC   STREET: 51 University Street
CC   CITY: Seattle
CC   STATE: WA
CC   COUNTRY: USA
CC   ZIP: 98101
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE: Floppy disk
CC   COMPUTER: Apple Macintosh
CC   OPERATING SYSTEM: Apple Operating System 7.1
CC   SOFTWARE: Microsoft Word for Apple, Version 5.1a
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: US/08/446,922
CC   FILING DATE:
CC   CLASSIFICATION: 435
CC   PRIOR APPLICATION DATA:
CC   APPLICATION NUMBER: USSN 08/107,353
CC   FILING DATE: 08-13-93
CC   CLASSIFICATION: 435
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME: Perkins, Patricia A
CC   REGISTRATION NUMBER: 34,693
CC   REFERENCE/DOCKET NUMBER: 1003-A
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE: (206)587-0430
CC   TELEFAX: (206)233-0644
CC   INFORMATION FOR SEQ ID NO: 4:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH: 261 amino acids
CC   TYPE: amino acid
CC   TOPOLOGY: linear
CC   MOLECULE TYPE: protein
SQ   SEQUENCE 261 AA; 29273 MW; 363115 CN;

Query Match
Best Local Similarity   7.2%;   Score 165; DB 1; Length 261;
Matches   44; Conservative   47; Mismatches   44; Indels   14; Gaps   12;

Db   122 IAAHVISEASSKTTSVLQW-AEKGYYTMSNNLVTLENGKQLTVKRQGLYIYAQVTFCSN 180
          | |::||:.:| |:| ::| ||:|| ||| ||||:::|::||:| :|:::||:|| :
Qy   169 INAASIPSGSHKVT-LSSWYHDRGWAKISN-M-TLSNGK-LRVNQDGFYLYANICF-RH 223

Db   181 REASSQAP--FIASL-CL-KSPGRF--ERILLRAANTHS-SAKPCGQ-QSIHLGGVFELQ 232
          :|:|| | : | :: | :|:::| | | ::|| ||| :
Qy   224 HETSGSVPTDYIQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLR 283

Db   233 PGASVFVNTDPSQVSHGTGFTSFGLLKL 261
          :| ::|:||:.:: | ||::.:|:
Qy   284 AGEEISIQVSNPSLLDPDQDATYFGAFKV 312

RESULT 10
ID   US-07-940-605A-2   STANDARD;   PRT;   261 AA.

```
XX
AC   xxxxxx
XX
DT
DE   Sequence 2, Application US/07940605A
XX
CC   Patent No. 5540926
CC   GENERAL INFORMATION:
CC      APPLICANT: ARUFFO, ALEJANDRO
CC      APPLICANT: HOLLENBAUGH, DIANE
CC      APPLICANT: LEDBETTER, JEFFREY A.
CC      TITLE OF INVENTION: SOLUBLE LIGANDS FOR CD40
CC      NUMBER OF SEQUENCES: 15
CC      CORRESPONDENCE ADDRESS:
CC         ADDRESSEE: Pennie & Edmonds
CC         STREET: 1155 Avenue of the Americas
CC         CITY: New York
CC         STATE: New York
CC         COUNTRY: U.S.A.
CC         ZIP: 10036-2711
CC      COMPUTER READABLE FORM:
CC         MEDIUM TYPE: Floppy disk
CC         COMPUTER: IBM PC compatible
CC         OPERATING SYSTEM: PC-DOS/MS-DOS
CC         SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/07/940,605A
CC         FILING DATE: 04-SEP-1992
CC         CLASSIFICATION: 530
CC      ATTORNEY/AGENT INFORMATION:
CC         NAME: Misrock, S. Leslie
CC         REGISTRATION NUMBER: 18,872
CC         REFERENCE/DOCKET NUMBER: 5624-184
CC      TELECOMMUNICATION INFORMATION:
CC         TELEPHONE: 212 790-9090
CC         TELEFAX: 212 869-8864/9741
CC         TELEX: 66141 PENNIE
CC      INFORMATION FOR SEQ ID NO: 2:
CC         SEQUENCE CHARACTERISTICS:
CC         LENGTH: 261 amino acids
CC         TYPE: amino acid
CC         TOPOLOGY: linear
CC      MOLECULE TYPE: protein
SQ   SEQUENCE 261 AA; 29273 MW; 363115 CN;

Query Match         7.2%;  Score 165;  DB 1;  Length 261;
Best Local Similarity 29.5%;  Pred. No. 4.71e-05;
Matches   44;  Conservative  47;  Mismatches 44;  Indels 14;  Gaps 12;

Db  122 IAAHVISEASSKTTSVLQW-AEKGYYTMSNNLVTLENGKQLTVKRQGLYIYAQVTFCSN 180
         | |:||||| :| : :||  :| ||| ||||||||||||||||||||| :||  :
Qy  169 INAASIPSGSHKVT-LSSWYHDRGWAKISN-M-TLSNGK-LRVNQDGFYYLYANICF-RH 223

Db  181 REASSQAP--FIASL-CL-KSPGRF--ERILLRAANTHS-SAKPCGQ-QSIHLGGVFELQ 232
         :|:|: |   : :|: : :|::::    ||::|::: ||  :|:||:
Qy  224 HETSGSVPTDYLQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLR 283

Db  233 PGASVFVNVTDPSQVSHGTGFTSFGLLKL 261
         :|: |||: ||:|| |:| ||| :|:
Qy  284 AGEEISIQVSNPSLLDPDQDATYFGAFKV 312

RESULT   11
ID   US-08-184-422-8    STANDARD;    PRT;    261 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 8, Application US/08184422
XX
```

```
CC   Sequence 8, Application US/08184422
CC   Patent No. 5565321
CC   GENERAL INFORMATION:
CC      APPLICANT: ARMITAGE, RICHARD
CC      APPLICANT: DAVISON, BARRY
CC      APPLICANT: FANSLOW, WILLIAM
CC      APPLICANT: RENSHAW, BLAIR
CC      APPLICANT: SPRIGGS, MELANIE
CC      APPLICANT: WIDMER, MICHAEL
CC      TITLE OF INVENTION: DETECTION AND TREATMENT OF MUTATIONS
CC      TITLE OF INVENTION: IN A CD40 LIGAND GENE
CC      NUMBER OF SEQUENCES: 16
CC      CORRESPONDENCE ADDRESS:
CC         ADDRESSEE: IMMUNEX CORPORATION
CC         STREET: 51 UNIVERSITY STREET
CC         CITY: SEATTLE
CC         STATE: WASHINGTON
CC         COUNTRY: USA
CC         ZIP: 98101
CC      COMPUTER READABLE FORM:
CC         MEDIUM TYPE: Floppy disk
CC         COMPUTER: Apple Macintosh
CC         OPERATING SYSTEM: Apple Operating System 7.1
CC         SOFTWARE: MS Word for Apple 5.1, Version a
CC      CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/08/184,422
CC         FILING DATE:
CC         CLASSIFICATION: 800
CC      PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER: 08/009,258
CC         FILING DATE: 01/22/93
CC      ATTORNEY/AGENT INFORMATION:
CC         NAME: PERKINS, PATRICIA ANNE
CC         REGISTRATION NUMBER: 34,693
CC         REFERENCE/DOCKET NUMBER: 2810-A
CC      TELECOMMUNICATION INFORMATION:
CC         TELEPHONE: 2065870430
CC         TELEFAX: 2065870606
CC      INFORMATION FOR SEQ ID NO: 8:
CC         SEQUENCE CHARACTERISTICS:
CC         LENGTH: 261 amino acids
CC         TYPE: amino acid
CC         TOPOLOGY: linear
CC      MOLECULE TYPE: protein
SQ   SEQUENCE 261 AA; 29273 MW; 363115 CN;

Query Match         7.2%;  Score 165;  DB 1;  Length 261;
Best Local Similarity 29.5%;  Pred. No. 4.71e-05;
Matches   44;  Conservative  47;  Mismatches 44;  Indels 14;  Gaps 12;

Db  122 IAAHVISEASSKTTSVLQW-AEKGYYTMSNNLVTLENGKQLTVKRQGLYIYAQVTFCSN 180
         | |:||||| :| : :||  :| ||| ||||||||||||||||||||| :||: :
Qy  169 INAASIPSGSHKVT-LSSWYHDRGWAKISN-M-TLSNGK-LRVNQDGFYYLYANICF-RH 223

Db  181 REASSQAP--FIASL-CL-KSPGRF--ERILLRAANTHS-SAKPCGQ-QSIHLGGVFELQ 232
         :|:|: |   : :|: : :|:::::    ||::|::: ||  :|:||:
Qy  224 HETSGSVPTDYLQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLR 283

Db  233 PGASVFVNVTDPSQVSHGTGFTSFGLLKL 261
         :|: |||: ||:|| |:| ||| :|:
Qy  284 AGEEISIQVSNPSLLDPDQDATYFGAFKV 312

RESULT   12
ID   PCT-US93-10034-4    STANDARD;    PRT;    261 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 4, Application PC/TUS9310034
XX
CC   Sequence 4, Application PC/TUS9310034
```

```
CC   GENERAL INFORMATION:
CC     APPLICANT: Spriggs, Melanie
CC     APPLICANT: Srinivasan, Subhashini
CC     TITLE OF INVENTION: Methods of Preparing Soluble, Oligomeric
CC     TITLE OF INVENTION: Proteins
CC     NUMBER OF SEQUENCES: 8
CC     CORRESPONDENCE ADDRESS:
CC     ADDRESSEE: Immunex Corporation
CC     STREET: 51 University Street
CC     CITY: Seattle
CC     STATE: WA
CC     COUNTRY: USA
CC     ZIP: 98101
CC     COMPUTER READABLE FORM:
CC     MEDIUM TYPE: Floppy disk
CC     COMPUTER: IBM PC compatible
CC     OPERATING SYSTEM: PC-DOS/MS-DOS
CC     SOFTWARE: PatentIn Release #1.0, Version #1.25
CC     CURRENT APPLICATION DATA:
CC     APPLICATION NUMBER: PCT/US93/10034
CC     FILING DATE:
CC     CLASSIFICATION:
CC     ATTORNEY/AGENT INFORMATION:
CC     NAME: Perkins, Patricia A
CC     REGISTRATION NUMBER: 34,693
CC     REFERENCE/DOCKET NUMBER: 1003
CC     TELECOMMUNICATION INFORMATION:
CC     TELEPHONE: (206)587-0430
CC     TELEFAX: (206)233-0644
CC   INFORMATION FOR SEQ ID NO: 4:
CC     SEQUENCE CHARACTERISTICS:
CC     LENGTH: 261 amino acids
CC     TYPE: amino acid
CC     TOPOLOGY: linear
CC     MOLECULE TYPE: protein
SQ   SEQUENCE   261 AA; 29273 MW; 363115 CN;

Query Match       7.2%;   Score 165; DB 2; Length 261;
Best Local Similarity 29.5%;   Pred. No. 4.7le-05;
Matches   44;   Conservative  47; Mismatches 44; Indels 14; Gaps 12;

Db  122 IAAHVISEASSKTTSVLQW-AEKGYYTMSNNLVTLENGKQLTVKRQGLYYIYAQVTFCSN 180
        | | :| :!!:!!!! : :!| !!! ||| !!|::!|::!!:| :: :
Qy  169 INAASIPSGSHKVT-LSSWYHDRGWAKISN-M-TLSNGK-LRVNQDGFYLYANICF-RH 223

Db  181 REASSQAP--FIASL-CL-KSPGRF--ERILLRAANTHS-SAKPCGQ-QSIHLGGVFELQ 232
        :!:!| : : |::: : |:::::!! : |::::::|! !:::: :||:::!| ! :|:
Qy  224 HETSGGSVPTDYLQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLR 283

Db  233 PGASVFVNVTDPSQVSHGTGFTSFGLLKL 261
        :| :|:!!| :!| ::!:
Qy  284 AGEEISIQVSNPSLLDPDQDATYFGAFKV 312

RESULT   13
ID   US-08-446-922-11     STANDARD;     PRT;     273 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 11, Application US/08446922
XX
CC   Sequence 11, Application US/08446922
CC   Patent No. 5716805
CC   GENERAL INFORMATION:
CC     APPLICANT: Spriggs, Melanie
CC     APPLICANT: Srinivasan, Subhashini
CC     TITLE OF INVENTION: Methods of Preparing Soluble, Oligomeric
CC     TITLE OF INVENTION: Proteins
CC     NUMBER OF SEQUENCES: 11
CC     CORRESPONDENCE ADDRESS:
CC     ADDRESSEE: Immunex Corporation
```

```
CC     STREET: 51 University Street
CC     CITY: Seattle
CC     STATE: WA
CC     COUNTRY: USA
CC     ZIP: 98101
CC     COMPUTER READABLE FORM:
CC     MEDIUM TYPE: Floppy disk
CC     COMPUTER: Apple Macintosh
CC     OPERATING SYSTEM: Apple Operating System 7.1
CC     SOFTWARE: Microsoft Word for Apple, Version 5.1a
CC     CURRENT APPLICATION DATA:
CC     APPLICATION NUMBER: US/08/446,922
CC     FILING DATE:
CC     CLASSIFICATION: 435
CC     PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER: USSN 08/107,353
CC     FILING DATE: 08-13-93
CC     CLASSIFICATION: 435
CC     ATTORNEY/AGENT INFORMATION:
CC     NAME: Perkins, Patricia A
CC     REGISTRATION NUMBER: 34,693
CC     REFERENCE/DOCKET NUMBER: 1003-A
CC     TELECOMMUNICATION INFORMATION:
CC     TELEPHONE: (206)587-0430
CC     TELEFAX: (206)233-0644
CC   INFORMATION FOR SEQ ID NO: 11:
CC     SEQUENCE CHARACTERISTICS:
CC     LENGTH: 273 amino acids
CC     TYPE: amino acid
CC     TOPOLOGY: linear
CC     MOLECULE TYPE: protein
SQ   SEQUENCE   273 AA; 30821 MW; 395120 CN;

Query Match       7.2%;   Score 165; DB 1; Length 273;
Best Local Similarity 29.5%;   Pred. No. 4.7le-05;
Matches   44;   Conservative  47; Mismatches 44; Indels 14; Gaps 12;

Db  134 IAAHVISEASSKTTSVLQW-AEKGYYTMSNNLVTLENGKQLTVKRQGLYYIYAQVTFCSN 192
        | | :| :!!:!!!! : :!| !!! ||| !!|::!|::!!:| :: :
Qy  169 INAASIPSGSHKVT-LSSWYHDRGWAKISN-M-TLSNGK-LRVNQDGFYLYANICF-RH 223

Db  193 REASSQAP--FIASL-CL-KSPGRF--ERILLRAANTHS-SAKPCGQ-QSIHLGGVFELQ 244
        :!:!| : : |::: : |:::::!! : |::::::|! !:::: :||:::!| ! :|:
Qy  224 HETSGGSVPTDYLQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLR 283

Db  245 PGASVFVNVTDPSQVSHGTGFTSFGLLKL 273
        :| :|:!!| :!| ::!:
Qy  284 AGEEISIQVSNPSLLDPDQDATYFGAFKV 312

RESULT   14
ID   PCT-US93-02475-9     STANDARD;     PRT;     157 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 9, Application PC/TUS9302475
XX
CC   Sequence 9, Application PC/TUS9302475
CC   GENERAL INFORMATION:
CC     APPLICANT: Wisnieski, Bernadine J.
CC     TITLE OF INVENTION: Tumor Necrosis Factor with Modified
CC     TITLE OF INVENTION: Ion Channel
CC     NUMBER OF SEQUENCES: 13
CC     CORRESPONDENCE ADDRESS:
CC     ADDRESSEE: Donald G. Lewis
CC     STREET: 8328 Regents Road #1E
CC     CITY: San Diego
CC     STATE: California
CC     COUNTRY: USA
CC     ZIP: 92122
CC     COMPUTER READABLE FORM:
```

```
CC    MEDIUM TYPE: Diskette, 3.5 inch, 1.44 M storage
CC    COMPUTER: VE System 386
CC    OPERATING SYSTEM: MS-DOS 5
CC    SOFTWARE: Word Perfect
CC    CURRENT APPLICATION DATA:
CC    APPLICATION NUMBER: PCT/US93/02475
CC    FILING DATE: 19930412
CC    CLASSIFICATION:
CC    PRIOR APPLICATION DATA:
CC    APPLICATION NUMBER: US 07/852,625
CC    FILING DATE: 12 March 1992
CC    ATTORNEY/AGENT INFORMATION:
CC    NAME: Donald G. Lewis
CC    REGISTRATION NUMBER: 28636
CC    REFERENCE/DOCKET NUMBER: BJW-2
CC    TELECOMMUNICATION INFORMATION:
CC    TELEPHONE: (619) 554-2421
CC    TELEFAX: (619) 554-6312
SQ    INFORMATION FOR SEQ ID NO: 9:
CC    SEQUENCE CHARACTERISTICS:
CC    LENGTH: 157 amino acids
CC    TYPE: AMINO ACIDS
CC    TOPOLOGY: linear
CC    MOLECULE TYPE: protein
CC    FEATURE:
CC    NAME/KEY: Tumor Necrosis Factor (murine)
CC    OTHER INFORMATION: A blank residue designated by
CC    OTHER INFORMATION: "Xaa" is inserted after residue No. 70 of murine
CC    OTHER INFORMATION: TNF and the sequence numbering is augmented by 1
CC    OTHER INFORMATION: starting with residue No. 71 in order to maximize
CC    OTHER INFORMATION: the sequence homology with human TNF.
CC    PUBLICATION INFORMATION:
CC    AUTHORS: Caput, D., Beutler, B. Hartog, K.
CC    AUTHORS: Thayer, R., Brown-Shimer, S. and
CC    AUTHORS: Cerami, A.
CC    TITLE: Identification of a Common Nucleotide
CC    TITLE: Sequence in the 3'-Untranslated Region of mRNA
CC    TITLE: Molecules Specifying Inflammatory Mediators.
CC    JOURNAL: Proc. National Academy of Science,
CC    JOURNAL: U.S.A.
CC    VOLUME: 83
CC    PAGES: 1670-1674
CC    DATE: 1986
CC    RELEVANT RESIDUES IN SEQ ID NO: 9: 1-157 (includes
CC    RELEVANT RESIDUES IN SEQ ID NO: one blank)
SQ    SEQUENCE 157 AA; 17304 MW; 137297 CN;

Query Match
Best Local Similarity 25.6%;   Pred. No. 8.20e-05;   Length 157;
Matches 33;   Conservative 34;   Mismatches 57;   Indels 5;   Gaps 5;

Db   28  WLSQRANALLANGMDLKDNQLVVPADGLYLVYSQVLFKGQGCPXDVVLLTHTVSRFAISY  87
             | :|| :||:| | ||| |: | ::|:|: ||: :|: ::  : :
Qy  186  WYHDRGWAKISN-MTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYVVKT 244

Db   88  QEKV-NLLSAVKSPCPKDTPEGAELKPWYEPIYLGGVFQLEKGDQLSAEVNLPKYLDFAE 146
             :|:: .|| :: | :| :|| |||| :: ||||| :|::|| | :
Qy  245  SIKIPSSHNLMKGGSTKNWSGNSEFH-FY-SINVGGFFKLRAGEEISIQVSNPSLLD-PD 301

Db  147  SGQVYFGVI 155
             ||| :||
Qy  302  QDATYFGAF 310
```

RESULT 15
ID   US-07-940-605A-4   STANDARD;   PRT;   151 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 4, Application US/07940605A
XX

```
CC    Sequence 4, Application US/07940605A
CC    Patent No. 5540926
CC    GENERAL INFORMATION:
CC    APPLICANT: ARUFFO, ALEJANDRO
CC    APPLICANT: HOLLENBAUGH, DIANE
CC    APPLICANT: LEDBETTER, JEFFREY A.
CC    TITLE OF INVENTION: SOLUBLE LIGANDS FOR CD40
CC    NUMBER OF SEQUENCES: 15
CC    CORRESPONDENCE ADDRESS:
CC    ADDRESSEE: Pennie & Edmonds
CC    STREET: 1155 Avenue of the Americas
CC    CITY: New York
CC    STATE: New York
CC    COUNTRY: U.S.A.
CC    ZIP: 10036-2711
CC    COMPUTER READABLE FORM:
CC    MEDIUM TYPE: Floppy disk
CC    COMPUTER: IBM PC compatible
CC    OPERATING SYSTEM: PC-DOS/MS-DOS
CC    SOFTWARE: PatentIn Release #1.0, Version #1.25
CC    CURRENT APPLICATION DATA:
CC    APPLICATION NUMBER: US/07/940,605A
CC    FILING DATE: 04-SEP-1992
CC    CLASSIFICATION: 530
CC    ATTORNEY/AGENT INFORMATION:
CC    NAME: Misrock, S. Leslie
CC    REGISTRATION NUMBER: 18,872
CC    REFERENCE/DOCKET NUMBER: 5624-184
CC    TELECOMMUNICATION INFORMATION:
CC    TELEPHONE: 212 790-9090
CC    TELEFAX: 212 869-8864/9741
CC    TELEX: 66141 PENNIE
CC    INFORMATION FOR SEQ ID NO: 4:
CC    SEQUENCE CHARACTERISTICS:
CC    LENGTH: 151 amino acids
CC    TYPE: amino acid
CC    TOPOLOGY: linear
CC    MOLECULE TYPE: protein
SQ    SEQUENCE 151 AA; 16904 MW; 127489 CN;

Query Match          6.7%;   Score 154;   DB 1;   Length 151;
Best Local Similarity 30.8%;   Pred. No. 3.57e-04;
Matches 45;   Conservative 45;   Mismatches 38;   Indels 18;   Gaps 16;

Db    4  QRGDEDPQIAAHV-VSEAN-SNAA-SV-LQ-W-AKKGYYTMKSNLVMLENGKQLTVKREG  57
             || :|| ||::|:::  ::::::  | :|  :: : || ::|| | :| | |::::|
Qy  155  QRGKPEAQPFAHLTINAASIPSGSHKVTLSSWYHDRGW-AKISNMT-LSNGK-LRVNQDG 211

Db   58  LYYYYTQVTFCSNREPSSQRP--FI-VGLWL-KPSI--GSERILLKAANTHSSSQLCE-Q 110
             :|| |:: | :||  | : :| | | ::|| ||:: | :::::| |: :|:|:: : -|:
Qy  212  FYYLYANICF-RHHETSGSVPTDYLQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFH 270

Db  111  -QSVHLGGYFELQAGASVFVNVTEAS 135
             |:::||| |:| |
Qy  271  FYSINVGGFFKLRAGEEISIQVSNPS 296
```

Search completed: Fri Nov 13 17:19:44 1998
Job time : 29 secs.

geck

```
***********************************************************
***********************************************************
       _    _   _____  _____   _____   _   _
      | |  | | |_   _||  ___| |  _  | | | | |
      | |  | |   | |  |___  | | | | | | |_| |
      | |  | |   | |   ___| | | |_| | |  _  |
      |_|  |_|   |_|  |_____| |_____| |_| |_|
                                                    (TM)
***********************************************************
***********************************************************

     Release 3.1A John F. Collins, Biocomputing Research Unit.
     Copyright (c) 1993-1998 University of Edinburgh, U.K.
           Distribution rights by Oxford Molecular Ltd

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:     Fri Nov 13 17:15:32 1998;  MasPar time 9.22 Seconds
                                        859.359 Million cell updates/sec
Tabular output not generated.

Title:          >US-08-989-362-2
Description:    (1-316) from US08989362.pep
Perfect Score:  2294
Sequence:       1 MRRASRDYGKYLRSSEEMGS........LLDPDQDATYFGAFKVQDID 316

Scoring table:  PAM 150
                Gap 11

Searched:       69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:       swiss-prot35
                1:swiss1

Statistics:     Mean 47.976;  Variance 83.538;  scale 0.574

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                                  SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	326	14.2	281	1	TRAI_HUMAN	TNF-RELATED APOPTOSIS	3.46e-44
2	282	12.3	291	1	TRAI_MOUSE	TNF-RELATED APOPTOSIS	7.11e-35
3	174	7.6	278	1	FASL_RAT	FAS ANTIGEN LIGAND.	1.89e-13
4	175	7.6	281	1	FASL_HUMAN	FAS ANTIGEN LIGAND (AP	1.23e-13
5	172	7.5	279	1	FASL_MOUSE	FAS ANTIGEN LIGAND.	4.40e-13
6	165	7.2	261	1	CD4L_HUMAN	CD40 LIGAND (CD40-L) (	8.25e-12
7	162	7.1	261	1	CD4L_BOVIN	CD40 LIGAND (TNF-RELAT	2.86e-11
8	160	7.0	235	1	TNFA_MOUSE	TUMOR NECROSIS FACTOR	6.51e-11
9	155	6.8	234	1	TNFA_CAVPO	TUMOR NECROSIS FACTOR	5.01e-10
10	152	6.6	235	1	TNFA_RABIT	TUMOR NECROSIS FACTOR	1.68e-09
11	150	6.5	235	1	TNFA_RAT	TUMOR NECROSIS FACTOR	3.76e-09
12	150	6.5	260	1	CD4L_MOUSE	CD40 LIGAND (TNF-RELAT	3.76e-09
13	145	6.3	235	1	TNFA_PERLE	TUMOR NECROSIS FACTOR	2.74e-08
14	143	6.2	234	1	TNFA_HORSE	TUMOR NECROSIS FACTOR	6.01e-08
15	136	5.9	306	1	TNFC_MOUSE	LYMPHOTOXIN-BETA (LT-B	9.04e-07
16	134	5.8	193	1	TNFA_CAPHI	TUMOR NECROSIS FACTOR	1.94e-06
17	132	5.8	233	1	TNFA_CANFA	TUMOR NECROSIS FACTOR	4.13e-06
18	128	5.6	233	1	TNFA_FELCA	TUMOR NECROSIS FACTOR	1.84e-05
19	127	5.5	229	1	TNFA_CEREL	TUMOR NECROSIS FACTOR	2.67e-05
20	127	5.5	233	1	TNFA_MACMU	TUMOR NECROSIS FACTOR	2.67e-05
21	126	5.5	233	1	TNFA_HUMAN	TUMOR NECROSIS FACTOR	3.86e-05
22	123	5.4	233	1	TNFA_PAPSP	TUMOR NECROSIS FACTOR	1.16e-04
23	123	5.4	233	1	TNFA_BOVIN	TUMOR NECROSIS FACTOR	1.16e-04
24	125	5.4	234	1	TNFA_SHEEP	TUMOR NECROSIS FACTOR	5.57e-05
25	122	5.3	205	1	TNFB_HUMAN	LYMPHOTOXIN-ALPHA PREC	1.66e-04
26	114	5.0	232	1	TNFA_PIG	TUMOR NECROSIS FACTOR	2.83e-03
27	111	4.8	244	1	TNFC_HUMAN	LYMPHOTOXIN-BETA (LT-B	7.93e-03
28	107	4.7	202	1	TNFB_MOUSE	LYMPHOTOXIN-ALPHA PREC	3.05e-02
29	106	4.6	345	1	CEBB_HUMAN	CCAAT/ENHANCER BINDING	4.25e-02
30	106	4.6	963	1	CHS2_YEAST	CHITIN SYNTHASE 2 (EC	4.25e-02
31	103	4.5	197	1	TNFB_RABIT	LYMPHOTOXIN-ALPHA PREC	1.13e-01
32	104	4.5	202	1	TNFB_RAT	LYMPHOTOXIN-ALPHA PREC	8.20e-02
33	101	4.4	657	1	SIM2_MOUSE	SINGLE-MINDED HOMOLOG	2.16e-01
34	99	4.3	395	1	RRPP_PI2H	RNA POLYMERASE ALPHA S	4.06e-01
35	99	4.3	395	1	RRPP_PI2HT	RNA POLYMERASE ALPHA S	4.06e-01
36	98	4.3	450	1	YK22_YEAST	HYPOTHETICAL 46.9 KD P	5.55e-01
37	98	4.3	520	1	WASP_MOUSE	WISKOTT-ALDRICH SYNDRO	5.55e-01
38	98	4.3	975	1	CDP_CANFA	CCAAT DISPLACEMENT PRO	5.55e-01
39	96	4.2	147	1	RNL4_PIG	RIBONUCLEASE PL3 PRECU	1.03e+00
40	97	4.2	542	1	ESTS_DROVI	ESTERASE S PRECURSOR (	7.57e-01
41	97	4.2	544	1	ESTP_DROME	ESTERASE P PRECURSOR (	7.57e-01
42	97	4.2	807	1	AFSK_STRGR	SERINE/THREONINE PROTE	7.57e-01
43	95	4.1	204	1	TNFB_BOVIN	LYMPHOTOXIN-ALPHA PREC	1.40e+00
44	94	4.1	1528	1	SPAA_STRDO	CELL SURFACE ANTIGEN I	1.89e+00
45	94	4.1	1805	1	HMW2_MYCGE	CYTADHERENCE HIGH MOLE	1.89e+00

```
                                 ALIGNMENTS

RESULT    1

ID   TRAI_HUMAN      STANDARD;      PRT;    281 AA.
AC   P50591;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   TNF-RELATED APOPTOSIS INDUCING LIGAND (TRAIL PROTEIN) (APO-2 LIGAND).
GN   TRAIL.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96111955.
RA   WILEY S.R., SCHOOLEY K., SMOLAK P.J., DIN W.S., HUANG C.-P.,
RA   NICHOLL J.K., SUTHERLAND G.R., DAVIS-SMITH T., RAUCH C., SMITH C.A.,
RA   GOODWIN R.G.;
RL   IMMUNITY 3:673-682(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RA   PITTI R.M., MARSTERS S.A., RUPPERT S., DONAHUE C.J., MOORE A.,
RA   ASHKENAZI A.;
RL   SUBMITTED (MAY-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: INDUCES APOPTOSIS.
CC   -!- SUBUNIT: HOMOTRIMER (POTENTIAL).
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (POTENTIAL).
CC   -!- TISSUE SPECIFICITY: WIDESPREAD; MOST PREDOMINANT IN SPLEEN, LUNG
CC       AND PROSTATE.
CC   -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR   EMBL; U37518; G1149558; -.
DR   EMBL; U57059; G1336208; -.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   CYTOKINE; TRANSMEMBRANE; SIGNAL-ANCHOR; APOPTOSIS.
FT   DOMAIN        1     17       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     18     38       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT   DOMAIN       39    281       EXTRACELLULAR (POTENTIAL).
SQ   SEQUENCE    281 AA;    32509 MW;  4D8OD8B1 CRC32;

Query Match           14.2%;    Score 326;   DB 1;  Length 281;
Best Local Similarity  35.9%;    Pred. No. 3.46e-44;
Matches   47;  Conservative   33;  Mismatches  49;  Indels   2;  Gaps   2;

Db  151 INSWESSRGHSFLSNLHLRNGELVIHEKGFYYIYSQTYFRFQEEIKENTKNDKQMVQYI 210
            :||: | | : ::|:|| :: ||:|  ||||| :|: |||:||  |::|:  ||:|: |::|
Qy  183 LSSWYHDR-GWAKISNMTLSNGKLRVNQDGFYLIYANICFRHETSGSVPTDYLQLMVYV 241
```

```
Db  211 YKYT-SYPDPILLMKSARNSCWSKDAEYGLYSIYQGGIFELKENDRIFVSVTNEHLIDMD 269
        |-|   ||||:   |||||  || |||||||:: :||| ||:| |:|
Qy  242 VKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDPD 301

Db  270 HEASFFGAFLV 280
        ::|:||||||
Qy  302 QDATYFGAFKV 312

RESULT  2
ID   TRAI_MOUSE    STANDARD;      PRT;     291 AA.
AC   P50592;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   TNF-RELATED APOPTOSIS INDUCING LIGAND (TRAIL PROTEIN).
GN   TRAIL.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96111955.
RA   WILEY S.R., SCHOOLEY K., SMOLAK P.J., DIN W.S., HUANG C.-P.,
RA   NICHOLL J.K., SUTHERLAND G.R., DAVIS-SMITH T., RAUCH C., SMITH C.A.,
RA   GOODWIN R.G.;
RL   IMMUNITY 3:673-682(1995).
CC   -!- FUNCTION: INDUCES APOPTOSIS.
CC   -!- SUBUNIT: HOMOTRIMER (POTENTIAL).
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (POTENTIAL).
CC   -!- TISSUE SPECIFICITY: WIDESPREAD.
CC   -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR   EMBL; U37522; G1149560; -.
DR   MGD; MGI:107414; TRAIL.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   CYTOKINE; TRANSMEMBRANE; SIGNAL-ANCHOR; APOPTOSIS.
FT   DOMAIN        1     17       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     18     38       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT   DOMAIN       39    291       EXTRACELLULAR (POTENTIAL).
FT   CARBOHYD     52     52       POTENTIAL.
SQ   SEQUENCE   291 AA;  33477 MW;  F9906FBE CRC32;

Query Match             12.3%;               Score 282;  DB 1;  Length 291;
Best Local Similarity   35.6%;               Pred. No. 7.11e-35;
Matches   48;  Conservative   37;  Mismatches   42;  Indels    8;  Gaps    5;

Db  155 IESWESSRKGHSFLNHVLFRNGELVIEQEGLYYISQTYFRFQEAEDASKMVSKDKVRTK 214
        || ||  :  :: |:|||  |||:|| |||:|:  || :: :  :: :
Qy  183 LSSWYHDR-GWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHET-SGS--VPTDYL--- 235

Db  215 QLVQYIYKYT-SYPDPIVLMKSARNSCWSRDAEYGLYSIYQGGLFELKNDRIFSVTNE 273
        :|| |  ||   ||||  |||||||||||  ||||||||||  || |:|||||  :|
Qy  236 QLMYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNP 295

Db  274 HLMDLDQEASFFGAF 288
        ||  |:|  |:||||
Qy  296 SLLDPDQDATYFGAF 310

RESULT  3
ID   FASL_RAT      STANDARD;      PRT;     278 AA.
AC   P36940;
DT   01-JUN-1994 (REL. 29, CREATED)
DT   01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   FAS ANTIGEN LIGAND.
GN   APT1LG1 OR FASL.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94084792.
RA   SUDA T., TAKAHASHI T., GOLSTEIN P., NAGATA S.;
RL   CELL 75:1169-1178(1993).
CC   -!- FUNCTION: CYTOKINE THAT BINDS TO FAS ANTIGEN, A RECEPTOR THAT
CC       TRANSDUCES THE APOPTOTIC SIGNAL INTO CELLS. MAY BE INVOLVED IN
CC       CYTOTOXYC T CELL MEDIATED APOPTOSIS AND IN T CELL DEVELOPMENT.
CC       FAS-ANTIGEN MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE INDUCTION OF
CC       PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED SUICIDE OF MATURE
CC       T CELLS, OR BOTH.
CC   -!- SUBUNIT: HOMOTRIMER (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MAY BE RELEASED
CC       INTO THE EXTRACELLULAR FLUID, PROBABLY BY CLEAVAGE FORM THE CELL
CC       SURFACE.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ACTIVATED SPLENOCYTES AND
CC       THYMOCYTES. MODERATE OR WEAK EXPRESSION FOUND IN SMALL INTESTINES,
CC       KIDNEY AND LUNG.
CC   -!- INDUCTION: BY PMA/IONOMYCIN AND CONCAVALIN/INTERLEUKIN-2.
CC   -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR   EMBL; U03470; G440179; -.
DR   HSSP; P19999; 1CLG.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   CYTOKINE; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR; APOPTOSIS.
FT   DOMAIN        1     77       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     78     99       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT   DOMAIN      100    278       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN        4     69       PRO-RICH.
FT   DOMAIN       45     58       POLY-PRO.
FT   DISULFID    199    230       BY SIMILARITY.
FT   CARBOHYD    116    116       POTENTIAL.
FT   CARBOHYD    247    247       POTENTIAL.
FT   CARBOHYD    257    257       POTENTIAL.
FT   CARBOHYD    116    116       POTENTIAL.
FT   CARBOHYD    247    247       POTENTIAL.
FT   CARBOHYD    257    257       POTENTIAL.
SQ   SEQUENCE   278 AA;  31140 MW;  6DC17725 CRC32;

Query Match             7.6%;                Score 174;  DB 1;  Length 278;
Best Local Similarity   26.7%;               Pred. No. 1.89e-13;
Matches   36;  Conservative   38;  Mismatches   53;  Indels    8;  Gaps    6;

Db  152 SRSIPLE-WEDTYGTALISGVKYKKGGLVINEAGLYFVYSKVVFRGQ-SCNSQP---LSH 206
        :|:: |  ||| |||:|| :| || |::|:||| :|  || |:||     |||||    ||
Qy  178 SHKVTLSSWYHDRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQL 237

Db  207 KVYM-R-NFKIYPGDLVLMEE-KKLNYCTTGQIWAHSSYLGAVFNLTVADHLYVNISQLSL 263
        |  | | ||| |  ::|:||:  ||||:|||:  ||| :|::|  |:::| :|::|  |
Qy  238 MYYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSL 297

Db  264 INFEESKTFFGLYKL 278
        ::| ::|  ||| |
Qy  298 LDPDQDATYFGAFKV 312

RESULT  4
ID   FASL_HUMAN    STANDARD;      PRT;     281 AA.
AC   P48023;
DT   01-FEB-1996 (REL. 33, CREATED)
DT   01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   FAS ANTIGEN LIGAND (APOPTOSIS ANTIGEN LIGAND) (APTL).
GN   APT1LG1 OR FASL.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95105731.
RA   ALDERSON M.;
RL   J. EXP. MED. 181:71-77(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
```

US-08-989-362-2.rsp

377-7

```
RX   MEDLINE; 95127560.
RA   TAKAHASHI T., TANAKA M., INAZAWA J., ABE T., SUDA T., NAGATA S.;
RL   INT. IMMUNOL. 6:1567-1574(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   SCHAETZLEIN C.E., POEHLMANN R., PHILIPPSEN P., EIBEL H.;
RL   SUBMITTED (JUN-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   MITA E., HAYASHI N., IIO S., TAKEHARA T., HIJIOKA T., KASAHARA A.,
RA   FUSAMOTO H., KAMADA T.;
RL   CELL. MOL. BIOL. RES. 204:468-474(1994).
RN   [5]
RP   SEQUENCE FROM N.A.
RA   WILKINSON J.;
RL   SUBMITTED (MAY-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: CYTOKINE THAT BINDS TO FAS ANTIGEN, A RECEPTOR THAT
CC       TRANSDUCES THE APOPTOTIC SIGNAL INTO CELLS. MAY BE INVOLVED IN
CC       CYTOTOXYC T CELL MEDIATED APOPTOSIS AND IN T CELL DEVELOPMENT.
CC       FAS-ANTIGEN MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE INDUCTION OF
CC       PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED SUICIDE OF MATURE
CC       T CELLS, OR BOTH.
CC   -!- SUBUNIT: HOMOTRIMER (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MAY BE RELEASED
CC       INTO THE EXTRACELLULAR FLUID, PROBABLY BY CLEAVAGE FROM THE CELL
CC       SURFACE.
CC   -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR   EMBL; X89102; G887456; -.
DR   EMBL; U08137; G624628; -.
DR   EMBL; U11821; G595431; -.
DR   EMBL; D38122; G1369902; -.
DR   EMBL; 296050; E320286; -.
MIM; 134638; -.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   CYTOKINE; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR; APOPTOSIS;
KW   DISEASE MUTATION.
FT   DOMAIN        1     80       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     81    102       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT   DOMAIN      103    281       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN        4     70       PRO-RICH.
FT   DOMAIN       45     65       POLY-PRO.
FT   DISULFID    202    233       BY SIMILARITY.
FT   CARBOHYD    184    184       POTENTIAL.
FT   CARBOHYD    250    250       POTENTIAL.
FT   CARBOHYD    260    260       POTENTIAL.
SQ   SEQUENCE   281 AA;  31485 MW;  714BAA90 CRC32;
```

Query Match   7.68%;   DB 1; Length 281;
Best Local Similarity 23.9%; Score 175; Pred. No. 1.23e-13;
Matches 44; Conservative 55; Mismatches 75; Indels 10; Gaps 8;

```
Db   107 LQKELAELR-ESTSQMHTASSLE-KQIGHPSPPEKKELRKVAHLTGKSNSRSMPLEWED 164
         ||   :|  |  ::| ::|  :| : :| |: :|:|  :   :||   ::| :::
Qy   130 LQHIVGPQRFSGAPAMMEGSWLDVAQRGKPEAQPFAHLTNAASIPSGSHKVTLS-SWYH 188

Db   165 TYGIVLLSGVKYKKGGLVINETGLYFVYSKVYFRGQ-SCNNLP---LSHKVYM-RNS-KY 218
         :: |: :|||||||||  | | :|||||::|| ||  |: | |   :|:|| | ||: ||
Qy   189 DRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHETSGSVPTDYLQLMVYVVKTSIKI 248

Db   219 PQDL-VMMEGKMMSYCTTGQMWARSSYLGAVFNLTSADHLYVNVSELSLVNFEESQTFFG 277
         |::| : |: :: :     :|  :::   | |:  |:  || |||  :|:| ::|  :||
Qy   249 PSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDPDQDATYFG 308

Db   278 LYKL 281
         :|:
Qy   309 AFKV 312
```

RESULT 5
ID   FASL_MOUSE   STANDARD;   PRT;   279 AA.
AC   P41047;
DT   01-FEB-1995 (REL. 31, CREATED)

```
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   FAS ANTIGEN LIGAND.
GN   APT1LG1 OR FASL OR GLD.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94185175.
RA   TAKAHASHI T., TANAKA M., BRANNAN C.I., JENKINS N.A., COPELAND N.G.,
RA   SUDA T., NAGATA S.;
RL   CELL 76:969-976(1994).
RN   [2]
RP   SEQUENCE FROM N.A., AND 3D-STRUCTURE MODELLING.
RC   STRAIN=C57BL/6;
RX   MEDLINE; 95388076.
RA   PEITSCH M.J., TSCHOPP J.J.;
RL   MOL. IMMUNOL. 32:761-772(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95196085.
RA   LYNCH D.H., WATSON M.L., ALDERSON M.R., BAUM P.R., MILLER R.E.,
RA   TOUGH T., GIBSON M., DAVIS-SMITH T., SMITH C.A., HUNTER K.;
RL   IMMUNITY 1:131-136(1994).
RN   [4]
RP   CHARACTERIZATION OF VARIANT GLD.
RX   MEDLINE; 96091792.
RA   HAHNE M., PEITSCH M.C., IRMLER M., SCHROETER M., LOWIN B.,
RA   ROUSSEAU M., BRON C., RENNO T., FRENCH L., TSCHOPP J.;
RL   INT. IMMUNOL. 7:1381-1386(1995).
CC   -!- FUNCTION: CYTOKINE THAT BINDS TO FAS ANTIGEN, A RECEPTOR THAT
CC       TRANSDUCES THE APOPTOTIC SIGNAL INTO CELLS. MAY BE INVOLVED IN
CC       CYTOTOXYC T CELL MEDIATED APOPTOSIS AND IN T CELL DEVELOPMENT
CC       FAS-ANTIGEN MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE INDUCTION OF
CC       PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED SUICIDE OF MATURE
CC       T CELLS, OR BOTH.
CC   -!- SUBUNIT: HOMOTRIMER (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MAY BE RELEASED
CC       INTO THE EXTRACELLULAR FLUID, PROBABLY BY CLEAVAGE FROM THE CELL
CC       SURFACE.
CC   -!- DISEASE: A DEFICIENCY IN THIS PROTEIN IS THE CAUSE OF GENERALIZED
CC       LYMPHOPROLIFERATION DISEASE (GLD), AN AUTOSOMAL RECESSIVE DISEASE
CC       RESPONSIBLE FOR LYMPHADENOPATHY AND AUTOANTIBODY PRODUCTION.
CC   -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR   EMBL; U06948; G473565; -.
DR   EMBL; U10984; G511222; -.
DR   EMBL; S76752; G913760; -.
DR   MGD; MGI:99255; FASL.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   CYTOKINE; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR; APOPTOSIS;
KW   DISEASE MUTATION.
FT   DOMAIN        1     78       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     79    100       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT   DOMAIN      101    279       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN        4     69       PRO-RICH.
FT   DOMAIN       45     51       POLY-PRO.
FT   DISULFID    200    231       BY SIMILARITY.
FT   VARIANT     273    273       F -> L (IN GLD; ABOLISH BIDING OF FASL TO
FT                                ITS RECEPTOR).
FT   CARBOHYD    117    117       POTENTIAL.
FT   CARBOHYD    182    182       POTENTIAL.
FT   CARBOHYD    248    248       POTENTIAL.
FT   CARBOHYD    258    258       POTENTIAL.
SQ   SEQUENCE   279 AA;  31442 MW;  AFD84D4B CRC32;
```

Query Match   7.5%;   DB 1; Length 279;
Best Local Similarity 25.9%; Score 172; Pred. No. 4.40e-13;
Matches 35; Conservative 39; Mismatches 53; Indels 8; Gaps 7;

```
Db   153 SRSIPLE-WEDTYGTALISGVKYKKGGLVINETGLYFVYSKVYFRGQ-SCNNQPLN--H- 207
         |: :| |  | ||   ||| :||:::  | |: ||:||:|::|| ||  || |:::  |:
```

Qy   178 SHKVTLSSWYHDRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQL 237
Db   208 KVYM-RNS-KYPEDLVLMEEKRL-NYCTTGQIWAHSSYLGAVFNLTSADHLYVNISQLSL 264
           :||:.|-:|-|| .|| |- .:|:.|| .:|:| ::||:|:.|| :.:||QVSNPSL
Qy   238 MVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEHFYSINVGGFFKLRAGEEISIQVSNPSL 297

Db   265 INFEESKTFFGLYKL 279
           :: :||. |:|| :|:
Qy   298 LDPDQDATYFGAFKV 312

RESULT   6
ID   CD4L_HUMAN   STANDARD;   PRT;   261 AA.
AC   P29965;
DT   01-APR-1993 (REL. 25, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   CD40 LIGAND (CD40-L) (TNF-RELATED ACTIVATION PROTEIN) (TRAP) (T CELL
DE   ANTIGEN GP39) (CD154 ANTIGEN).
CN   CD40LG OR CD40L OR TRAP.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93076854.
RA   GRAF D., KORTHAEUER U., MAGES H.W., SENGER G., KROCZEK R.A.;
RL   EUR. J. IMMUNOL. 22:3191-3194(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93145330.
RA   HOLLENBAUGH D., GROSMAIRE L., KULLAS C., CHALUPNY J.,
RA   BRAESCH-ANDERSEN S., NOELLE R., STAMENKOVIC I., LEDBETTER J.,
RA   ARUFFO A.;
RL   EMBO J. 11:4313-4321(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93145330.
RA   SPRIGGS M.K., ARMITAGE R.J., STOCKBINE L., CLIFFORD K.N.,
RA   MACDUFF B.M., SATO T.A., MALISZEWSKI C.R., FANSLOW W.C.;
RL   CELL 72:291-300(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93094757.
RA   SPRIGGS M.K., ARMITAGE R.J., STROCKBINE L., CLIFFORD K.N.,
RA   MACDUFF B.M., SATO T.A., MALISZEWSKI C.R., FANSLOW W.C.;
RL   J. EXP. MED. 176:1543-1550(1992).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93138085.
RA   GAUCHAT J.F.M., AUBRY J., MAZZEI G.J., LIFE P., JOMOTTE T., ELSON G.,
RA   BONNEFOY J.Y.;
RL   FEBS LETT. 315:259-266(1993).
RN   [6]
RP   SEQUENCE FROM N.A.
RA   SHIMADZU M., TERASAKI H., NINOMIYA R., SHIMIZU S., NUNOI H.,
RA   MATSUDA I.;
RL   SUBMITTED (FEB-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 116-261.
RX   MEDLINE; 96131874.
RA   KARPUSAS M., HSU Y.-M., WANG J.-H., THOMPSON J., LEDERMAN S.,
RA   CHESS L., THOMAS D.;
RL   STRUCTURE 3:1031-1039(1995).
RN   [8]
RP   VARIANTS HIGM1 ARG-36 AND GLU-140.
RX   MEDLINE; 93156839.
RA   KORTHAEUER U., GRAF D., MAGES H.W., BRIERE F., PADAYACHEE M.,
RA   MALCOLM S., UGAZIO A.G., NOTARANGELO L.D., LEVINSKY R.J.,
RA   KROCZEK R.A.;
RL   NATURE 361:539-541(1993).
RN   [9]
RP   VARIANT HIGM1 GLU-123.

RX   MEDLINE; 93156840.
RA   DISANTO J.P., BONNEFOY J.Y., GAUCHAT J.F., FISCHER A.,
RA   DE SAINT BASILE G.;
RL   NATURE 361:541-543(1993).
RN   [10]
RP   VARIANTS HIGM1 ARG-128; GLY-129 AND PRO-235.
RX   MEDLINE; 93145330.
RA   ARUFFO A., FARRINGTON M., HOLLENBAUGH D., LI X., MILATOVICH A.,
RA   NONOYAMA S., BAJORATH J., GROSMAIRE L.S., STENKAMP R., NEUBAUER M.,
RA   ROBERTS R.L., NOELLE R.J., LEDBETTER J.A., FRANCKE U., OCHS H.D.;
RL   CELL 72:291-300(1993).
RN   [11]
RP   VARIANTS HIGM1 PRO-155; ASP-211 AND VAL-227.
RX   MEDLINE; 93174270.
RA   ALLEN R.C., ARMITAGE R.J., CONLEY M.E., ROSENBLATT H., JENKINS N.A.,
RA   COPELAND N.G., BEDELL M.A., EDELHOFF S., DISTECHE C.M.,
RA   SIMONEAUX D.K., FANSLOW W.C., BELMONT J., SPRIGGS M.K.;
RL   SCIENCE 259:990-993(1993).
RN   [12]
RP   VARIANTS HIGM1 ALA-126; ARG-140 AND GLU-144.
RX   MEDLINE; 95233438.
RA   MACCHI P., VILLA A., STRINA D., SACCO M.G., MORALI F., BRUGNONI D.,
RA   GILIANI S., MANTUANO E., FASTH A., ANDERSSON B., ZEGERS B.J.M.,
RA   CAVAGNI G., REZNICK I., LEVY J., ZAN-BAR I., PORAT Y., AIRO P.,
RA   PLEBANI A., VEZZONI P., NOTARANGELO L.D.;
RL   AM. J. HUM. GENET. 56:898-906(1995).
RN   [13]
RP   VARIANTS HIGM1 ARG-36; CYS-140; SER-231; MET-254 AND GLY-227 DEL.
RX   MEDLINE; 97295077.
RA   NONOYAMA S., SHIMADZU M., TORU H., SEYAMA K., NUNOI H., NEUBAUER M.,
RA   YATA J.-I., OCH H.D.;
RL   HUM. GENET. 99:624-627(1997).
CC   -!- FUNCTION: MEDIATES B-CELL PROLIFERATION IN THE ABSENCE OF CO-
CC   STIMULUS AS WELL AS IGE PRODUCTION IN THE PRESENCE OF IL-4.
CC   INVOLVED IN IMMUNOGLOBULIN CLASS SWITCHING.
CC   -!- SUBUNIT: HOMOTRIMER.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC   EXTRACELLULAR SOLUBLE FORM.
CC   -!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED ON ACTIVATED CD4+
CC   T-LYMPHOCYTES.
CC   -!- DISEASE: DEFECTS IN CD40LG ARE THE CAUSE OF AN X-LINKED
CC   IMMUNODEFICIENCY WITH HYPER-IGM (HIGM1), AN IMMUNOGLOBULIN ISOTYPE
CC   SWITCH DEFECT CHARACTERIZED BY ELEVATED CONCENTRATIONS OF SERUM
CC   IGM AND DECREASED AMOUNTS OF ALL OTHER ISOTYPES. AFFECTED MALES
CC   PRESENT AT AN EARLY AGE (USUALLY WITHIN THE FIRST YEAR OF LIFE)
CC   RECURRENT BACTERIAL AND OPPORTUNISTIC INFECTIONS, INCLUDING
CC   PNEUMOCYSTIS CARINII PNEUMONIA AND INTRACTABLE DIARRHEA DUE TO
CC   CRYPTOSPORIDIUM INFECTION. DESPITE SUBSTITUTION TREATMENT WITH
CC   INTRAVENOUS IMMUNOGLOBULIN, THE OVERALL PROGNOSIS IS RATHER POOR,
CC   WITH A DEATH RATE OF ABOUT 10% BEFORE ADOLESCENCE.
CC   -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC   -!- DATABASE: NAME=CD40Lbase; NOTE=European CD40L defect database;
CC   WWW="http://www.expasy.ch/www/cd40lbase.html";
CC   FTP="ftp.expasy.ch/databases/cd40lbase".
DR   EMBL; X68550; G37270; -.
DR   EMBL; Z15017; G38484; -.
DR   EMBL; X67878; G38412; -.
DR   EMBL; L07414; G180124; -.
DR   EMBL; D31797; G1518170; JOINED.
DR   EMBL; D31794; G1518170; JOINED.
DR   EMBL; D31795; G1518170; JOINED.
DR   EMBL; D31796; G1518170; JOINED.
DR   PIR; S25684; S25684.
DR   PIR; S26694; S26694.
DR   PIR; S28017; S28017.
DR   PIR; S28852; S28852.
DR   PIR; JH0793; JH0793.
DR   PDB; 1ALY; 17-SEP-97.
DR   MIM; 308230; -.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS55049; TNF_2; 1.
KW   CYTOKINE; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR; 3D-STRUCTURE;

```
KW   DISEASE MUTATION.
FT   DOMAIN       1     22       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    23     46       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT   DOMAIN      47    261       EXTRACELLULAR (POTENTIAL).
FT   DISULFID   178    218       POTENTIAL.
FT   CARBOHYD   240    240       POTENTIAL.
FT   VARIANT     36     36       M -> R (IN H1GM1).
FT   VARIANT    123    123       A -> E (IN H1GM1).
FT   VARIANT    126    126       V -> A (IN H1GM1).
FT   VARIANT    128    129       SE -> RG (IN H1GM1).
FT   VARIANT    140    140       W -> C (IN H1GM1).
FT   VARIANT    140    140       W -> G (IN H1GM1).
FT   VARIANT    140    140       W -> R (IN H1GM1).
FT   VARIANT    144    144       G -> E (IN H1GM1).
FT   VARIANT    155    155       L -> P (IN H1GM1).
FT   VARIANT    211    211       T -> D (IN H1GM1).
FT   VARIANT    227    227       G -> V (IN H1GM1).
FT   VARIANT    227    227       MISSING (IN H1GM1).
FT   VARIANT    231    231       L -> S (IN H1GM1).
FT   VARIANT    235    235       A -> P (IN H1GM1).
FT   VARIANT    254    254       T -> M (IN H1GM1).
SQ   SEQUENCE   261 AA;  29273 MW;  DC2AD21F CRC32;

Query Match              7.2%;   Score 165; DB 1;  Length 261;
Best Local Similarity   29.5%;   Pred. No. 8.25e-12;
Matches   44;  Conservative   47;  Mismatches   44;  Indels   14;  Gaps   12;

Db  122 IAAHVISEASSKTTSVLQW-AEKGYYTMSNNLVTLENGKQLTVKRQGLYYIYAQVTFCSN 180
            | :  |::: ||:|| :::::::|  :: |: |:: ||::| |::::||||::|: |: |
Qy  169 INAASIPSGSHKVT-LSSWYHDRGWAKISN-M-TLSNGK-LRVNQDGFYLYANICF-RH 223

Db  181 REASSQAP--FIASL-CL-KSPGRF--ERILLRAANTHS-SAKPCGQ-QSIHLGGVFELQ 232
             :|:|  :|:  ||  :|  |: :||||:| : |:|:::::||  | |: ||:|| :| |
Qy  224 HETSGSVPTDYLQLMVYVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLR 283

Db  233 PGASVFVNVTDPSQVSHGTGFTSFGLLKL 261
            :|| | ::|:|  |  :|| | || ::
Qy  284 AGEEISIQVSNPSLLDPDQDATYFGAFKV 312

RESULT   7
ID   CD4L_BOVIN    STANDARD;   PRT;   261 AA.
AC   P51749;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   CD40 LIGAND (TNF-RELATED ACTIVATION PROTEIN) (TRAP) (T CELL ANTIGEN
DE   GP39).
GN   CD40LG OR CD40L.
OS   BOS TAURUS (BOVINE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; ARTIODACTYLA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BLOOD;
RX   MEDLINE; 96006582.
RA   MERTENS B.E.L.C., MURIUKI M., GAIDULIS L.;
RL   IMMUNOGENETICS 42:430-431(1995).
CC   -!- FUNCTION: MEDIATES B-CELL PROLIFERATION IN THE ABSENCE OF CO-
CC       STIMULUS AS WELL AS IGE PRODUCTION IN THE PRESENCE OF IL-4.
CC       INVOLVED IN IMMUNOGLOBULIN CLASS SWITCHING (BY SIMILARITY).
CC   -!- SUBUNIT: HOMOTRIMER.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC       EXTRACELLULAR SOLUBLE FORM (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR   EMBL; Z48469; G732570; -.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   CYTOKINE; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT   DOMAIN       1     22       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    23     46       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT   DOMAIN      47    261       EXTRACELLULAR (POTENTIAL).
FT   DISULFID   178    218       POTENTIAL.
```

```
FT   CARBOHYD   240    240       POTENTIAL.
SQ   SEQUENCE   261 AA;  29242 MW;  649521FA CRC32;

Query Match              7.1%;   Pred. No. 2.86e-11;
Best Local Similarity   30.2%;   Score 162; DB 1;  Length 261;
Matches   45;  Conservative   40;  Mismatches   50;  Indels   14;  Gaps   13;

Db  122 IAAHVISEASSKTTSVLQW-APKGYYTLSNNLVTLENGKQLAVKRQGFYYIYTQVTFCSN 180
            | :  |::: ||:|| :::::::|  :: |: |:: ||::| |::::||||::|: |: |
Qy  169 INAASIPSGSHKVT-LSSWYHDRGWAKISN-M-TLSNGK-LRVNQDGFYLYANICF-RH 223

Db  181 RETLSQAP--FIASL-CL-KSP-S-GSERILLRAANTHSSSKPCG-Q-QSIHLGGVFELQ 232
             ::|:|  :|:  ||  :|  |: :||||:| : |:|:::::||  | |: ||:|| :| |
Qy  224 HETSGSVPTDYLQLMVYVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLR 283

Db  233 SGASVFVNVTDPSQVSHGTGFTSFGLLKL 261
            :|| | ::|:|  |  :|| | || ::
Qy  284 AGEEISIQVSNPSLLDPDQDATYFGAFKV 312

RESULT   8
ID   TNFA_MOUSE    STANDARD;   PRT;   235 AA.
AC   P06804;
DT   01-JAN-1988 (REL. 06, CREATED)
DT   01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN   TNFA OR TNF.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 88224564.
RA   SHIRAI T., SHIMIZU N., SHIOJIRI S., HORIGUCHI S., ITO H.;
RL   DNA 7:193-201(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 85298296.
RA   PENNICA D., HAYFLICK J.S., BRINGMAN T.S., PALLADINO M.A.,
RA   GOEDDEL D.V.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 82:6060-6064(1985).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86149365.
RA   FRANSEN L., MULLER R., MARMENOUT A., TAVERNIER J., VAN DER HEYDEN J.,
RA   KAWASHIMA E., CHOLLET A., TIZARD R., VAN HEUVERSWYN H., VAN VLIET A.,
RA   RUYSSCHAERT M.-R., FIERS W.;
RL   NUCLEIC ACIDS RES. 13:4417-4429(1985).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   CAPUT D., BEUTLER B., HARTOG K., THAYER R., BROWN-SHIMER S.,
RA   CERAMI A.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 83:1670-1674(1986).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87298639.
RA   SHAKHOV A.N., NEDOSPASOV S.A.;
RL   BIOORG. KHIM. 13:701-705(1987).
RN   [6]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 88067722.
RA   SEMON D., KAWASHIMA E., JONGENEEL C.V., SHAKHOV A.N., NEDOSPASOV S.A.;
RL   NUCLEIC ACIDS RES. 15:9083-9084(1987).
RN   [7]
RP   SEQUENCE OF 80-99.
RX   MEDLINE; 91097531.
RA   SHERRY B., JUC D.-M., ZENTELLA A., CERAMI A.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 173:1072-1078(1990).
RN   [8]
RP   SEQUENCE OF 70-87.
RX   MEDLINE; 89380231.
```

```
RA   CSEH K., BEUTLER B.;
RL   J. BIOL. CHEM. 264:16256-16260(1989).
RN   [9]
RP   IDENTIFICATION OF MEMBRANE-BOUND FORM.
RX   MEDLINE; 88165056.
RA   KRIEGLER M., PEREZ X., DEFAY K., ALBERT I., LU S.D.;
RL   CELL 53:45-53(1988).
CC   -!- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC       WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC       CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC       CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC       OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC       CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC       CONDITIONS.
CC   -!- SUBUNIT: HOMOTRIMER.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC       EXTRACELLULAR SOLUBLE FORM.
CC   -!- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC       PROTEOLYTIC PROCESSING.
CC   -!- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC       CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC       AND MALNUTRITION.
CC   -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR   EMBL; U06950; G49468; -.
DR   EMBL; M13049; G202083; -.
DR   EMBL; M11731; G202085; -.
DR   EMBL; Y00467; G54832; -.
DR   EMBL; X02611; G54845; -.
DR   EMBL; M20155; G202093; ALT_SEQ.
DR   EMBL; M38296; G202087; -.
DR   PIR; A23127; QWMSN.
DR   PIR; A22908; A22908.
DR   PIR; A25164; A25164.
DR   PIR; A27303; A27303.
DR   PIR; A34251; A34251.
DR   PIR; S03791; S03791.
DR   HSSP; P01375; 1TNF.
DR   MGD; MGI:104798; TNF.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT   PROPEP        1     79
FT   CHAIN        80    235       TUMOR NECROSIS FACTOR.
FT   TRANSMEM     36     56       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT   DISULFID    148    179       BY SIMILARITY.
FT   CARBOHYD     86     86
FT   CONFLICT    231    231       G -> R (IN REF. 3 AND 4).
SQ   SEQUENCE   235 AA;  25895 MW;  7FDE3C71 CRC32;

Query Match            7.0%;  Score 160;  DB 1;  Length 235;
Best Local Similarity  25.6%;  Pred. No. 6.51e-11;
Matches   33;  Conservative  34;  Mismatches 56;  Indels  6;  Gaps  6;

Db   107 WLSQRANALLANGMDLKDNQLVVPADGLYLVVSQVLFKGQGCPDYVLLTH-TVSRFAISY 165
         |  :|| :|::|:||:| |||:| | |:| :::|  :  |   :|:
Qy   186 WYHDRGWAKISN-MTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYVVKT 244

Db   166 QEKV-NLLSAVKSPCPKDTPEGAELKPWEYPIYLGGVFQLEKGDQLSAEVNLPKYLDFAE 224
         ::: :|: :|::  ::  : | ||| | :| |:| ||  ::: :|| :|
Qy   245 SIKIPSSHNLMKGGSTKNWSGNSEFH-FY-SINVGGFFKLRAGEEISIQVSNPSLLD-PD 301

Db   225 SGQVYFGVI 233
         ||| :
Qy   302 QDATYFGAF 310

RESULT 9
ID   TNFA_CAVPO     STANDARD;      PRT;    234 AA.
AC   P51435;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN   TNFA.
OS   CAVIA PORCELLUS (GUINEA PIG).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HARTLEY; TISSUE=LUNG;
RA   YUAN H.T., KELLY F.J., BINGLE C.D.;
RL   SUBMITTED (NOV-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DUNKIN-HARTLEY;
RA   WHITE A.M., YOSHIMURA T., SMITH A.W., WESTWICK J., WATSON M.L.;
RL   SUBMITTED (NOV-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC       WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC       CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC       CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC       OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC       CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC       CONDITIONS.
CC   -!- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC       EXTRACELLULAR SOLUBLE FORM (BY SIMILARITY).
CC   -!- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC       PROTEOLYTIC PROCESSING (BY SIMILARITY).
CC   -!- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC       CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC       AND MALNUTRITION.
CC   -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR   EMBL; U39839; G1066112; -.
DR   EMBL; U77036; G1679724; -.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT   PROPEP        1     79       BY SIMILARITY.
FT   CHAIN        80    234       TUMOR NECROSIS FACTOR.
FT   TRANSMEM     36     56       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT   DISULFID    147    178       BY SIMILARITY.
SQ   SEQUENCE   234 AA;  25793 MW;  00160783 CRC32;

Query Match            6.8%;  Score 155;  DB 1;  Length 234;
Best Local Similarity  31.8%;  Pred. No. 5.01e-10;
Matches   42;  Conservative  28;  Mismatches 50;  Indels 12;  Gaps  8;

Db   106 WLSKRANALLANGMGLSDNQLVVPSDGLYLIYSQVLFKGQ--GC-PS-YLLLTHTVSRL 160
         |  :  |: || | :|| | :  |:| |:||  |  ::  :|| :  :  |:
Qy   186 WYHDRGWAKISN-MTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYVVKT 244

Db   161 AVSYPEKVNLLSAIKSPCQKETPEGAERKPWYEPIYLGGVFQLQKGDRLSAEVNLPQYLD 220
         :  |::  |   |  :  | : : : | ||| | :| |:| ||  :
Qy   245 SIKIPSSHNLM---KGGSTKNWSGNSEFH-FY-SINVGGFFKLRAGEEISIQVSNPSLLD 299

Db   221 FADSGQIYFGVI 232
         :|| ||| :
Qy   300 -PDQDATYFGAF 310

RESULT 10
ID   TNFA_RABIT     STANDARD;      PRT;    235 AA.
AC   P04924;
DT   13-AUG-1987 (REL. 05, CREATED)
DT   13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
DT   01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE   TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN   TNFA.
OS   ORYCTOLAGUS CUNICULUS (RABBIT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; LAGOMORPHA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91065534.
RA   SHAKHOV A.N., KUPRASH D.V., AZIZOV M.M., JONGENEEL C.V.,
```

```
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=TESTIS;
RX   MEDLINE; 94040766.
RA   KWON J., CHUNG I.Y., BENVENISTE E.N.;
RL   GENE 132:227-236(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86219711.
RA   SHIRAI T., SHIMIZU N., HORIGUCHI S., ITO H.;
RL   AGRIC. BIOL. CHEM. 53:1733-1736(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92329007.
RA   ESTLER H.C., GREWE M., GAUSSLING R., PAVLOVIC M., DECKER K.;
RL   BIOL. CHEM. HOPPE-SEYLER 373:271-281(1992).
RN   [4]
RP   SEQUENCE OF 1-231 FROM N.A.
RC   TISSUE=TAIL;
RA   KIRISITS M.J., VARDIMON D., KUNZ H.W., GILL T.J. III;
RL   SUBMITTED (JUN-1993) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC       WITH A WIDE VARIETY OF FUNCTIONS. IT CAN CAUSE CYTOLYSIS OF
CC       CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC       CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC       OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC       CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC       CONDITIONS.
CC   -!- SUBUNIT: HOMOTRIMER.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC       EXTRACELLULAR SOLUBLE FORM.
CC   -!- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC       PROTEOLYTIC PROCESSING.
CC   -!- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC       CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC       AND MALNUTRITION.
CC   -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR   EMBL; X66539; G395370; -.
DR   EMBL; L00981; G205254; -.
DR   EMBL; D00475; G220921; -.
DR   EMBL; L19123; G310232; -.
DR   PIR; JU0029; JU0029.
DR   PIR; S21674; S21674.
DR   PIR; JN0868; JN0868.
DR   HSSP; P01375; 1TNF.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT   PROPEP        1     79
FT   CHAIN        80    235       TUMOR NECROSIS FACTOR.
FT   TRANSMEM     36     56       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT   DISULFID    148    179       BY SIMILARITY.
FT   CARBOHYD     86     86       POTENTIAL.
FT   CONFLICT     39     39       L -> P (IN REF. 3).
FT   CONFLICT    163    163       I -> T (IN REF. 3).
FT   CONFLICT    202    202       F -> S (IN REF. 3).
SQ   SEQUENCE   235 AA;  25806 MW;  C1DF18B7 CRC32;

Query Match             6.5%;  Score 150;  DB 1;  Length 235;
Best Local Similarity  25.6%;  Pred. No. 3.76e-09;
Matches   33;  Conservative  32;  Mismatches  58;  Indels  6;  Gaps  6;

Db        107  WLSQRANALLANGMDLKDNQLVVPADGLYLIYSQVLFKGQGCPDYVLLTH-TVSRFAISY 165
Qy        186  WYHDRGWAKISN-MTLSNGKLRVNQDGFYLYANICFRHHETSGGSVPTDYLQLMVYVVKT 244
Db        166  QEKV-SLLSAIKSPCPKDTPEGAELKPWYEPMYLGGVFQLEKGDLLSAEVNLPKYLDITE 224
Qy        245  SIKIPSSHNLMKGGSTKNWSGNSEEFH-FY-SINVGGFFKLRAGEEISIQVSNPSLLD-PD 301
Db        225  SGQVYFGVI 233
             |||  :
```

```
RA   NEDOSPASOV S.A.;
RL   GENE 95:215-221(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86219712.
RA   ITO H., SHIRAI T., YAMAMOTO S., AKIRA M., KAWAHARA S., TODD C.W.,
RA   WALLACE R.B.;
RL   DNA 5:157-165(1986).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86219711.
RA   ITO H., YAMAMOTO S., KURODA S., SAKAMOTO H., KAJIHARA J., KIYOTA T.,
RA   HAYASHI H., KATO M., SEKO M.;
RL   DNA 5:149-156(1986).
CC   -!- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC       WITH A WIDE VARIETY OF FUNCTIONS. IT CAN CAUSE CYTOLYSIS OF
CC       CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC       CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC       OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC       CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC       CONDITIONS.
CC   -!- SUBUNIT: HOMOTRIMER.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC       EXTRACELLULAR SOLUBLE FORM.
CC   -!- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC       PROTEOLYTIC PROCESSING.
CC   -!- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC       CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC       AND MALNUTRITION.
CC   -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR   EMBL; M12845; G165760; -.
DR   EMBL; M12846; G165753; -.
DR   EMBL; M60340; G165756; -.
DR   PIR; A25451; A25451.
DR   PIR; A25454; A25454.
DR   PIR; JS0727; JS0727.
DR   HSSP; P01375; 1TNF.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT   PROPEP        1     79
FT   CHAIN        80    235       TUMOR NECROSIS FACTOR.
FT   TRANSMEM     36     56       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT   DISULFID    148    179       BY SIMILARITY.
FT   CONFLICT     63     63       MISSING (IN REF. 3).
SQ   SEQUENCE   235 AA;  25816 MW;  1BCE5CD8 CRC32;

Query Match             6.6%;  Score 152;  DB 1;  Length 235;
Best Local Similarity  26.0%;  Pred. No. 1.68e-09;
Matches   33;  Conservative  33;  Mismatches  55;  Indels  6;  Gaps  6;

Db        107  WLSQRANALLANGMKLTDNQLVVPADGLYLIYSQVLFSGQGCRSYVLLTH-TVSRPAVSY 165
Qy        186  WYHDRGWAKISN-MTLSNGKLRVNQDGFYLYANICFRHHETSGGSVPTDYLQLMVYVVKT 244
Db        166  PNKV-NLLSAIKSPCHRETPEEAEPMAWYEPIYLGGVFQLEKGDRLSTEVNQPEYLDLAE 224
Qy        245  SIKIPSSHNLMKGGSTKNWSGNSE-FHFY-SINVGGFFKLRAGEEISIQVSNPSLLD-PD 301
Db        225  SGQVYFG 231
             |||
Qy        302  QDATYFG 308

RESULT   11
ID   TNFA_RAT       STANDARD;      PRT;     235 AA.
AC   P16599;
DT   01-AUG-1990 (REL. 15, CREATED)
DT   01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN   TNFA.
OS   RATTUS NORVEGICUS (RAT).
```

Qy   302 QDATYFGAF 310

```
RESULT    12
ID   CD4L_MOUSE    STANDARD;    PRT;    260 AA.
AC   P27548;
DT   01-AUG-1992 (REL. 23, CREATED)
DT   01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   CD40 LIGAND (TNF-RELATED ACTIVATION PROTEIN) (TRAP) (T CELL ANTIGEN
DE   GP39).
GN   CD40LG OR CD40L.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92244364.
RA   ARMITAGE R., FANSLOW W., SATO T.A., CLIFFORD K.N., STROCKBINE L.,
RA   MACDUFF B.M., ANDERSON D.M., GIMPEL S.D., DAVIS-SMITH T.,
RA   MALISZEWSKI C.R., CLARK E.A., SMITH C.A., GRABSTEIN K.H., COSMAN D.,
RA   SPRIGGS M.K.;
RL   NATURE 357:80-82(1992).
RN   [2]
RP   SIMILARITY TO THE TNF FAMILY.
RX   MEDLINE; 92310561.
RA   FARRAH T., SMITH C.A.;
RL   NATURE 358:26-26(1992).
RN   [3]
RP   3D-STRUCTURE MODELLING OF 115-260.
RX   MEDLINE; 93200072.
RA   PEITSCH M.C., JONGENEEL C.V.;
RL   INT. IMMUNOL. 5:233-238(1993).
CC   -!- FUNCTION: MEDIATES B-CELL PROLIFERATION IN THE ABSENCE OF CO-
CC       STIMULUS AS WELL AS IGE PRODUCTION IN THE PRESENCE OF IL-4.
CC       INVOLVED IN IMMUNOGLOBULIN CLASS SWITCHING.
CC   -!- SUBUNIT: HOMOTRIMER.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC       EXTRACELLULAR SOLUBLE FORM.
CC   -!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED ON ACTIVATED CD4+
CC       T-LYMPHOCYTES.
CC   -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR   EMBL; X65453; G50352; -.
DR   PIR; S21738; S21738.
DR   PDB; 1CDA; 31-OCT-93.
DR   MGD; MGI:88337; CD40L.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   CYTOKINE; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR; 3D-STRUCTURE.
FT   DOMAIN        1     22      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     23     46      SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT   DOMAIN       47    260      EXTRACELLULAR (POTENTIAL).
FT   DISULFID    177    217      POTENTIAL.
FT   CARBOHYD    239    239      POTENTIAL.
SQ   SEQUENCE   260 AA;  29396 MW;  420FF8C9 CRC32;

Query Match   Best Local Similarity   6.5%;   Score 150;  DB 1;  Length 260;
Best Local Similarity   29.6%;   Pred. No. 3.76e-09;
Matches   48;   Conservative   49;   Mismatches   47;   Indels   18;   Gaps   16;

Db   113 QRGDEDPQIAAHV-VSEAN-SNAA-SV-LQ-W-AKKGYYTMKSNLVMLENGKQLTVKREG 166
             |||     ::| |:  ::::  |     |:   ||:  |:    || |  |||
Qy   155 QRGKPEAQPFAHLTINAASIPSGSHKVTLSSWYHDRGW-AKISNMT-LSNGK-LRVNQDG 211

Db   167 LYYVYTQVTFCSNREPSSORP--FI-VGLWL-KPSI--GSERILLKAANTHSSSQLCE-Q 219
             :||:|  :  |  ||   :| ::  || | :|   |:|::::||:  :  :
Qy   212 FYYLYANICF-RHHETSGSVPTDYLQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFH 270

Db   220 -QSVHLGGVFELQAGASVFVNVTEASQVIHRVGPSSFGLLKL 260
          ||:    |:  |:  ::   :|  |:
Qy   271 FYSINVGGFFKLRAGEEISIQVSNPSLLDPDQDATYFGAFKV 312
```

```
RESULT    13
ID   TNFA_PERLE    STANDARD;    PRT;    235 AA.
AC   P36939;
DT   01-JUN-1994 (REL. 29, CREATED)
DT   01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT   01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE   TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN   TNFA.
OS   PEROMYSCUS LEUCOPUS (WHITE-FOOTED MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92218012.
RA   CREW M.D., FILIPOWSKY M.E.;
RL   IMMUNOGENETICS 35:351-353(1992).
CC   -!- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC       WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC       CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC       CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC       OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC       CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC       CONDITIONS.
CC   -!- SUBUNIT: HOMOTRIMER.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC       EXTRACELLULAR SOLUBLE FORM.
CC   -!- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC       PROTEOLYTIC PROCESSING.
CC   -!- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC       CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC       AND MALNUTRITION.
CC   -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR   EMBL; M59233; G202507; -.
DR   HSSP; P01375; 1TNF.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT   PROPEP        1     79      BY SIMILARITY.
FT   CHAIN        80    235      TUMOR NECROSIS FACTOR.
FT   TRANSMEM     36     56      SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT   DISULFID    148    179      BY SIMILARITY.
FT   CARBOHYD     86     86      POTENTIAL.
SQ   SEQUENCE   235 AA;  25822 MW;  0D9AF4E8 CRC32;

Query Match   Best Local Similarity   6.3%;   Score 145;  DB 1;  Length 235;
Best Local Similarity   26.3%;   Pred. No. 2.74e-08;
Matches   31;   Conservative   29;   Mismatches   53;   Indels   5;   Gaps   5;

Db   118 NGMDLKDNQLVIPADGLYLVYSQVLFKGQGCSSYVLLTH-TVSRFAVSYEDKV-NLLSAI 175
             ::  | :|  |||:   ||:|| |:||   |: :  | |  :: ::    |: |
Qy   196 SNMTLSNGKLRVNQDGFPYLYANICFRHHETSGSVPTDYLQLMVYVVKTSIKIPSSHNLM 255

Db   176 KSPCPKETPEGSELKPWYEPIYLGGVFQLEKGDRLSAEVNLPKYLDFAESGQVYFGVI 233
             | |  |  ::| |  ::::| ||| | ::|  :|:
Qy   256 KGGSTKNWSGNSEFH-FY-SINVGGFFKLRAGEEISIQVSNPSLLD-PDQDATYFGAF 310
```

```
RESULT    14
ID   TNFA_HORSE    STANDARD;    PRT;    234 AA.
AC   P29553;
DT   01-APR-1993 (REL. 25, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE   TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN   TNFA.
OS   EQUUS CABALLUS (HORSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PERISSODACTYLA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92084125.
RA   SU X., MORRIS D.D., MCGRAW R.A.;
RL   GENE 107:319-321(1991).
```

```
CC   -!- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC       WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC       CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC       CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC       OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC       CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC       CONDITIONS.
CC   -!- SUBUNIT: HOMOTRIMER.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC       EXTRACELLULAR SOLUBLE FORM.
CC   -!- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC       PROTEOLYTIC PROCESSING.
CC   -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR   EMBL; M64087; G164245; -.
DR   PIR; JQ1344; JQ1344.
DR   HSSP; P01375; 1TNF.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
     PROPEP        1     77       BY SIMILARITY.
     CHAIN        78    234       TUMOR NECROSIS FACTOR.
FT   TRANSMEM     36     56       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
     DISULFID    146    178       BY SIMILARITY.
SQ   SEQUENCE   234 AA;  25469 MW;  CD20BE6E CRC32;

Query Match             6.2%;
Best Local Similarity  27.7%;     Score 143;  DB 1;  Length 234;
                                  Pred. No. 6.01e-08;
Matches  36;  Conservative  33;  Mismatches 50;  Indels 11;  Gaps 7;

Db  105 WLSGRANALLANGVKLTDNQLVVPLDGLYLIYSQVLFKGQ---GCPSTHVL-LTHTISRL 160
Qy  186 WYHDRGWAKISN-MTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYYVKT 244
Db  161 AVSYPSKVNLLSAIKSPCHTESPEQAEAKPWYEPIYLGGVFQLEKGDQLSAEINQPNYLD 220
Qy  245 SIKIPSSHNLM---KGGSTKNWSGNSEFH-FY-SINVGGFFKLRAGEEISIQVSNPSLLD 299
Db  221 FAESGQYYFG 230
Qy  300 -PDQDATYFG 308

RESULT 15
ID   TNFC_MOUSE     STANDARD;      PRT;    306 AA.
AC   P41155;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
     LYMPHOTOXIN-BETA (LT-BETA) (TUMOR NECROSIS FACTOR C).
     LTB OR TNFC.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RA   POKHOLOK D.K., MAROULAKOU I.G., KUPRASH D.V., ALIMZHANOV M.B.,
RA   KOZLOV S.V., NOVOBRANTSEVA T.I., TURETSKAYA R.L., GREEN J.E.,
RA   NEDOSPASOV S.A.;
RL   SUBMITTED (JUL-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=SPLEEN;
RX   MEDLINE; 95088371.
RA   LAWTON P., NELSON J., TIZARD R., BROWNING J.L.;
RL   J. IMMUNOL. 154:239-246(1995).
CC   -!- FUNCTION: MAY PLAY A SPECIFIC ROLE IN IMMUNE REPONSE REGULATION.
CC       PROVIDES THE MEMBRANE ANCHOR FOR THE ATTACHMENT OF THE HETEROMERIC
CC       COMPLEX TO THE CELL SURFACE.
CC   -!- SUBUNIT: HETEROTRIMER OF EITHER TWO BETA AND ONE ALPHA SUBUNITS
CC       OR (LESS PREVALENT) TWO ALPHA AND ONE BETA SUBUNITS.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.

DR   EMBL; U12029; G515509; -.
DR   EMBL; U16984; G577432; -.
DR   EMBL; U16985; G577831; -.
DR   EMBL; U06950; G495466; -.
DR   MGD; MGI:104796; LTB.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT   DOMAIN        1     27       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     28     48       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT   DOMAIN       49    306       EXTRACELLULAR (POTENTIAL).
FT   CARBOHYD     98     98       POTENTIAL.
FT   CARBOHYD    284    284       POTENTIAL.
SQ   SEQUENCE   306 AA;  32328 MW;  7C9780D1 CRC32;

Query Match             5.9%;
Best Local Similarity  28.7%;     Score 136;  DB 1;  Length 306;
                                  Pred. No. 9.04e-07;
Matches  33;  Conservative  28;  Mismatches 47;  Indels 7;  Gaps 6;

Db  191 LALPQDGVYYLYCHVGYRGRTPAGRSRARSLTLRSALYRAGGAYGRGSPELLLEGAETV 250
Qy  205 LRVNQDGFYYLYANICFR-HHETSGSVPTDYLQLM--VYVVKTSIKIPSSHNLMKGGSTK 261
Db  251 TPVVDPIGYGSLWYTSVGFGGGLAQLRSGERVVVNISHPDMVDYRRGKTFFGAVMV 305
Qy  262 NWSGNS-EF-HF-Y-SINVGGFFKLRAGEEISIQVSNPSLLDPDQDATYFGAFKV 312

Search completed: Fri Nov 13 17:16:27 1998
Job time : 55 secs.
```

```
**************************************************

  M P S R C H

                                    (TM)
**************************************************

        Release 3.1A John F. Collins, Biocomputing Research Unit.
        Copyright (c) 1993-1998 University of Edinburgh, U.K.
                Distribution rights by Oxford Molecular Ltd

MPsrch_pp  protein - protein database search, using Smith-Waterman algorithm

Run on:          Fri Nov 13 17:16:46 1998;  MasPar time 19.37 Seconds
                                812.531 Million cell updates/sec
Tabular output not generated.

Title:           >US-08-989-362-2
Description:     (1-316) from US08989362.pep
Perfect Score:   2294
Sequence:        1 MRRASRDYGKYLRSSEEMGS........LLDPDQDATYFGAFKVQDID 316

Scoring table:   PAM 150
                 Gap 11

Searched:        165420 seqs, 49795644 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        sptrembl6
                 1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
                 5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
                 9:sp_phage 10:sp_plant 11:sp_rodent 12:sp_unclassified
                 13:sp_vertebrate 14:sp_virus

Statistics:      Mean 46.565;  Variance 93.226;  scale 0.499

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

                        SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	2294	100.0	316	11	O35235	TRANCE.	0.00e+00
2	2287	99.7	316	11	O35306	RANKL.	0.00e+00
3	1986	86.6	317	4	O14788	RANKL.	0.00e+00
4	1527	66.6	245	4	O14723	TRANCE (FRAGMENT).	5.51e-284
5	175	7.6	279	11	Q61217	FAS LIGAND.	1.58e-11
6	161	7.0	232	11	O35853	TUMOR NECROSIS FACTOR	3.97e-09
7	161	7.0	233	11	O35734	TUMOR NECROSIS FACTOR.	2.76e-09
8	152	6.6	156	11	Q62326	TUMOR NECROSIS FACTOR.	7.03e-08
9	148	6.5	240	4	O43557	TUMOR NECROSIS FACTOR	2.89e-07
10	142	6.2	234	6	Q28320	TNF-ALPHA.	2.35e-06
11	125	5.4	216	11	O70332	TUMOR NECROSIS FACTOR-	7.11e-04
12	123	5.4	233	6	O18779	TUMOR NECROSIS FACTOR	1.36e-03
13	122	5.3	157	4	O43647	TUMOR NECROSIS FACTOR	1.88e-03
14	120	5.2	1055	11	O54967	NON-RECEPTOR PROTEIN T	3.56e-03
15	106	4.6	348	6	O02755	CCAAT/ENHANCER-BINDING	2.64e-01
16	105	4.6	375	10	P93066	OLEOSIN-LIKE PROTEIN.	3.54e-01
17	106	4.6	642	2	Q55103	CHO-ORF2.	2.64e-01
18	104	4.5	1091	4	Q07912	TYROSINE-PROTEIN KINAS	4.75e-01
19	100	4.4	200	13	O73770	DGCR6 HOMOLOG.	1.50e+00
20	101	4.4	223	8	Q34914	CYTOCHROME B (FRAGMENT	1.13e+00
21	100	4.4	223	8	Q34861	CYTOCHROME B (FRAGMENT	1.50e+00
22	100	4.4	223	8	Q34885	CYTOCHROME B (FRAGMENT	1.50e+00
23	100	4.4	223	8	Q34932	CYTOCHROME B (FRAGMENT	1.50e+00
24	100	4.4	223	8	Q34933	CYTOCHROME B (FRAGMENT	1.50e+00
25	100	4.4	223	8	Q34871	CYTOCHROME B (FRAGMENT	1.50e+00
26	100	4.4	232	7	Q31553	MAJOR HISTOCOMPATIBILI	1.50e+00
27	100	4.4	263	13	O73896	B LOCUS L BETA CHAIN 2	1.50e+00
28	100	4.4	263	7	Q31410	MHC CLASS II B-LBII-BE	1.50e+00
29	100	4.4	266	7	Q31413	MHC CLASS II BETA CHAI	1.50e+00
30	101	4.4	473	11	Q63887	NONO.	1.13e+00
31	101	4.4	4436	11	O58659	4436AA LONG HYPOTHETIC	1.13e+00
32	98	4.3	520	11	Q61078	WISCOTT-ALDRICH SYNDRO	2.64e+00
33	99	4.3	657	11	O35391	SINGLE-MINDED 2 PROTEI	1.99e+00
34	99	4.3	663	5	O61917	MUTL HOMOLOG.	1.99e+00
35	99	4.3	1616	4	O15054	KIAA0346 (FRAGMENT).	1.99e+00
36	96	4.2	160	8	O21745	CYTOCHROME B (FRAGMENT	4.60e+00
37	97	4.2	223	8	Q34899	CYTOCHROME B (FRAGMENT	4.60e+00
38	96	4.2	223	8	Q34906	CYTOCHROME B (FRAGMENT	4.60e+00
39	96	4.2	223	8	Q34934	CYTOCHROME B (FRAGMENT	4.60e+00
40	96	4.2	223	8	Q34924	CYTOCHROME B (FRAGMENT	4.60e+00
41	97	4.2	1217	5	O17889	F54F12.1 (EC 3.1.3.48)	3.49e+00
42	97	4.2	1343	14	Q06635	TRANS-ACTING TRANSCRIP	3.49e+00
43	97	4.2	5071	5	P91905	RYANODINE RECEPTOR.	3.49e+00
44	97	4.2	5107	5	Q94279	PARTIAL CDS.	3.49e+00
45	96	4.2	5262	4	O14686	ALR.	4.60e+00

```
                                ALIGNMENTS

RESULT    1
ID   O35235            PRELIMINARY;        PRT;   316 AA.
AC   O35235;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   TRANCE.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97460112.
RA   WONG B.R., RHO J., ARRON J., ORLINICK J., CHAO M.,
RA   KALACHIKOV S., CAYANI E., BARTLETT F.S. III, FRANKEL W.N., LEE S.Y.,
RA   CHOI Y.;
RL   J. BIOL. CHEM. 272:25190-25194(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   LACEY D.L., TIMMS E., TAN H.-L., KELLEY M.J., DUNSTAN C.R.,
RA   BURGESS T., ELLIOTT R., COLOMBERO A., ELLIOTT G., SCULLY S., HSU H.,
RA   SULLIVAN J., HAWKINS N., DAVY E., CAPPARELLI C., ELI A., QIAN Y.-X.,
RA   KAUFMAN S., SAROSI I., SHALHOUB V., SENALDI G., GUO J., DELANEY J.,
RA   BOYLE W.J.;
RL   CELL 93:165-176(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BONE MARROW;
RA   YASUDA H., SHIMA N., NAKAGAWA N., YAMAGUCHI K., KINOSAKI M.,
RA   MOCHIZUKI S., TOMOYASU A., YANO K., GOTO M., MURAKAMI A., TSUDA E.,
RA   MORINAGA T., HIGASHIO K., UDAGAWA N., TAKAHASHI N., SUDA T.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 95:3597-3602(1998).
DR   EMBL; AF013170; G2411498; -.
DR   EMBL; AF053713; G3057148; -.
DR   EMBL; AB008426; D1026353; -.
DR   PFAM; PF00229; TNF.
SQ   SEQUENCE 316 AA;   34944 MW;  F76EC806 CRC32;

     Query Match        100.0%;   Score 2294;  DB 11;   Length 316;
     Best Local Similarity 100.0%;  Pred. No. 0.00e+00;
     Matches  316;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Db      1 MRRASRDYGKYLRSSEEMGSGGPGVPHEGPLHPAPSAPAGPPPAASRSMFLALLGLGLGQ 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy        1 MRRASRDYGKYLRSSEEMGSGPGVPHEGPLHPAPSAPAPAPPPAASRSMFLALLGLGLGQ 60

Db       61 VVCSIALFLYFRAQMDPNRISEDSTHCFYRILRLHENAGLQDSTLESEDTLPDSCRRMKQ 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       61 VVCSIALFLYFRAQMDPNRISEDSTHCFYRILRLHENAGLQDSTLESEDTLPDSCRRMKQ 120

Db      121 AFQGAVQKELQHIVGPQRFSGAPAMMEGSWLDVAQRGKPEAQPFAHLTINAASIPSGSHK 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      121 AFQGAVQKELQHIVGPQRFSGAPAMMEGSWLDVAQRGKPEAQPFAHLTINAASIPSGSHK 180

Db      181 VTLSSWYHDRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGGVPTDYLQLMVY 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      181 VTLSSWYHDRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGGVPTDYLQLMVY 240

Db      241 VVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDP 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      241 VVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDP 300

Db      301 DQDATYFGAFKVQDID 316
            ||||||||||||||||
Qy      301 DQDATYFGAFKVQDID 316
```

RESULT 2
```
ID   O35306      PRELIMINARY;       PRT;    316 AA.
AC   O35306;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE   RANKL.
GN   RANKL.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE: 98032977.
RA   ANDERSON D.M., MARASKOVSKY E., BILLINGSLEY W.L., DOUGALL W.C.,
RA   TOMETSKO M.E., ROUX E.R., TEEPE M.C., DUBOSE R.F., COSMAN D.,
RA   GALIBERT L.;
RL   NATURE 390:175-179(1997).
DR   EMBL; AF019048; G2612924; -.
DR   PFAM; PF00229; TNF.
SQ   SEQUENCE  316 AA;  35002 MW;  97DA4504 CRC32;
```

```
Query Match             99.7%;  Score 2287;  DB 11;  Length 316;
Best Local Similarity   99.7%;  Pred. No. 0.00e+00;
Matches  315;  Conservative   0;  Mismatches   0;  Indels   1;  Gaps   0;

Db        1 MRRASRDYGKYLRSSEEMGSGPGVPHEGPLHPAPSAPAPAPPPAASRSMFLALLGLGLGQ 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        1 MRRASRDYGKYLRSSEEMGSGPGVPHEGPLHPAPSAPAPAPPPAASRSMFLALLGLGLGQ 60

Db       61 VVCSIALFLYFRAQMDPNRISEDSTHCFYRILRLHENADLQDSTLESEDTLPDSCRRMKQ 120
            ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Qy       61 VVCSIALFLYFRAQMDPNRISEDSTHCFYRILRLHENAGLQDSTLESEDTLPDSCRRMKQ 120

Db      121 AFQGAVQKELQHIVGPQRFSGAPAMMEGSWLDVAQRGKPEAQPFAHLTINAASIPSGSHK 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      121 AFQGAVQKELQHIVGPQRFSGAPAMMEGSWLDVAQRGKPEAQPFAHLTINAASIPSGSHK 180

Db      181 VTLSSWYHDRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGGVPTDYLQLMVY 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      181 VTLSSWYHDRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGGVPTDYLQLMVY 240

Db      241 VVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDP 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      241 VVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDP 300

Db      301 DQDATYFGAFKVQDID 316
            ||||||||||||||||
Qy      301 DQDATYFGAFKVQDID 316
```

```
Qy        1 MRRASRDYGKYLRSSEEMGSGPGVPHEGPLHPAPSAPAPAPPPAASRSMFVALLGLGLGQ 60
```

RESULT 3
```
ID   O14788      PRELIMINARY;       PRT;    317 AA.
AC   O14788;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   RANKL.
GN   RANKL.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RX   MEDLINE: 98032977.
RA   ANDERSON D.M., MARASKOVSKY E., BILLINGSLEY W.L., DOUGALL W.C.,
RA   TOMETSKO M.E., ROUX E.R., TEEPE M.C., DUBOSE R.F., COSMAN D.,
RA   GALIBERT L.;
RL   NATURE 390:175-179(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   LACEY D.L., TIMMS E., TAN H.-L., KELLEY M.J., DUNSTAN C.R.,
RA   BURGESS T., ELLIOTT R., COLOMBERO A., ELLIOTT G., SCULLY S., HSU H.,
RA   SULLIVAN J., HAWKINS N., DAVY E., CAPPARELLI C., ELI A., QIAN Y.-X.,
RA   KAUFMAN S., SAROSI I., SHALHOUB V., SENALDI G., GUO J., DELANEY J.,
RA   BOYLE W.J.;
RL   CELL 93:165-176(1998).
DR   EMBL; AF019047; G2612922; -.
DR   EMBL; AF053712; G3057146; -.
DR   PFAM; PF00229; TNF.
SQ   SEQUENCE  317 AA;  35478 MW;  A58E6DC4 CRC32;
```

```
Query Match             86.6%;  Score 1986;  DB 4;  Length 317;
Best Local Similarity   84.6%;  Pred. No. 0.00e+00;
Matches  269;  Conservative  30;  Mismatches  16;  Indels   3;  Gaps   3;

Db        1 MRRASRDYTKYLRGSEEMGGGPGAPHEGPLH-APPPPAPHQPPAASRSMFVALLGLGLGQ 59
            |||||||| ||||| ||||||  || |||||| |  :::||  |||||||||||||||||
Qy        1 MRRASRDYGKYLRSSEEMGSGPGVPHEGPLHPAPSAPAPAPPPAASRSMFVALLGLGLGQ 60

Db       60 VVCSVALFFYFRAQMDPNRISEDGTHCIYRIRLRHENADFQDTTLESODTKLIPDSCRRI 119
            ||||:|||:||||||||||||||.|||:|||.|||||| ::||:|||.||.:|:|||||
Qy       61 VVCSIALFLYFRAQMDPNRISEDSTHCFYRILRLHENAGLQDSTLESEDT-L-PDSCRRM 118

Db      120 KQAFQGAVQKELQHIVGSQHIRAEKAMVDGSWLDLAKRSKLEAQPFAHLTINATDIPSGS 179
            ||||||||||||||||| : :  :||  :|||||:|:| .| |||||||||||| :|||||
Qy      119 KQAFQGAVQKELQHIVGPQRFSGAPAMMEGSWLDVAQRGKPEAQPFAHLTINAASIPSGS 178

Db      180 HKVSLSSWYHDRGWAKISNMTFSNGKLLVNQDGFYYLYANICFRHHETSGDLATEYLQLM 239
            |||:|||||||||||||||||:|||||:|||||||||||||||||||||||.:.:||||| 
Qy      179 HKVTLSSWYHDRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLM 238

Db      240 VYVTKTSIKIPSSHTLMKGGSTKYWSGNSEFHFYSINVGGFFKLRSGEEISIEVSNPSLL 299
            | |.|||||||||||:|||||||.|||||||||||||||||||||:||||||:|||||||
Qy      239 VVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLL 298

Db      300 DPDQDATYFGAFKVRDID 317
            |||||||||||||| |||
Qy      299 DPDQDATYFGAFKVQDID 316
```

RESULT 4
```
ID   O14723      PRELIMINARY;       PRT;    245 AA.
AC   O14723;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE   TRANCE (FRAGMENT).
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
```

RX   MEDLINE; 97460112.
RA   WONG B.R., RHO J., ARRON J., ROBINSON E., ORLINICK J., CHAO M.,
RA   KALACHIKOV S., CAYANI E., BARTLETT F.S. III, FRANKEL W.N., LEE S.Y.,
RA   CHOI Y.;
RL   J. BIOL. CHEM. 272:25190-25194(1997).
DR   EMBL; AF013171; G2411500; -.
DR   PFAM; PF00229; TNF.
FT   NON_TER   1   1
SQ   SEQUENCE   245 AA; 27804 MW; F7CD1ECE CRC32;

Query Match   66.6%; Score 1527; DB 4; Length 245;
Best Local Similarity   84.5%; Pred. No. 5.51e-284;
  Matches 207; Conservative   24; Mismatches   12; Indels   2; Gaps   2;

Db   1 QMDPNRISEDGTHCIYRILRLHENADFQDTTLESQDTKLIPDSCRRIKQAFQGAVQKELQ 60
              |||||||||:||::|||||||||||:||:|||:|| :|||||||:|||||||||||||||
Qy   74 QMDPNRISEDSTHCFYRILRLHENAGLQDSLESEDT-L-PDSCRRMKQAFQGAVQKELQ 131

Db   61 HIVGSQHIRAEKAMVDGSWLDLAKRSKLEAQPFAHLTINATDIPSGSHKVSLSSWYHDRG 120
              ||||:|::|:::::|:|||||:|:|:||||||||||||||::||||||||:|||||||||
Qy   132 HIVGPQRFSGAPAMMEGSWLDVAQRGKPEAQPFAHLTINAASIPSGSHKVTLSSWYHDRG 191

Db   121 WGKISNMTPSNGKLIVNQDGFYYLYANICFRHHETSGDLATEYLQLMVYVYKTSIKIPSS 180
              |:|||||| :||||:|||||||||||||||||||||| :::||||||||| |||||||||
Qy   192 WAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYVVKTSIKIPSS 251

Db   181 HTLMKGGSTKYWSGNSEFHFYSINVGGFFKLRSGEEISIEVSNPSLLDPDDATYFGAFK 240
              | |||||||| |||||||||||||||||||||::|||||:|||||||||||:|||||||||
Qy   252 HNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDPDQDATYFGAFK 311

Db   241 VRDID 245
              | ||| |
Qy   312 VQDID 316

RESULT 5
ID   Q61217
AC   Q61217;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   FAS LIGAND.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RP   STRAIN=BALB/C;
RA   FENNER M.H., SHIODA T., ISSELBACHER K.J.;
RL   SUBMITTED (MAY-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; U58995; G1389772; -.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PFAM; PF00229; TNF.
SQ   SEQUENCE   279 AA; 31340 MW; 9062DF08 CRC32;

Query Match   7.6%; Score 175; DB 11; Length 279;
Best Local Similarity   25.9%; Pred. No. 1.58e-11;
  Matches   35; Conservative   40; Mismatches   52; Indels   8; Gaps   7;

Db   153 SRSIPLE-WEDTYGTALISCVKYKKGGLVINEAGLYFVSKVYFRGQ-SCNNQPLN--H- 207
              :::::|| | : | ||| :| |::::| :|| :::||:::::|::: |
Qy   178 SHKVTLSSWYHDRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQL 237

Db   208 KVYM-RNS-KYPGDLVLMEEKRL-NYCTTGQIWAHSSYLGAVFNLTSADHLYVNISQLSL 264
              ||:|: |:|: : || :: || :: :|: || :: || :|| :: :: : :|| ||
Qy   238 MVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSL 297

Db   265 INFEESKTFGLIYKL 279
              :|: ::: |: :
Qy   298 LDPDQDATYFGAFKV 312

RESULT 6

ID   O35853
AC   O35853;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   TUMOR NECROSIS FACTOR ALPHA.
GN   TNFA.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A/J;
RX   MEDLINE; 97245744.
RA   IRAQI F., TEALE A.;
RL   IMMUNOGENETICS 45:459-461(1997).
DR   EMBL; U68414; G2304957; -.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PFAM; PF00229; TNF.
SQ   SEQUENCE   232 AA; 25513 MW; 9B23F06 CRC32;

Query Match   7.0%; Score 160; DB 11; Length 232;
Best Local Similarity   25.6%; Pred. No. 3.97e-09;
  Matches   33; Conservative   34; Mismatches   56; Indels   6; Gaps   6;

Db   104 WLSQRANALLANGMDLKDNQLVVPADGLILVYSQVLFKGQGCPDYVLLTH-TVSRFAISY 162
              | :: :: ||| : : :|||| |:::|:|:| ||:|| | ||| | ::: : || ||: :|
Qy   186 WYHDRGWAKISN-MTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYVVKT 244

Db   163 QEKV-NLLSAVKSPCPKDTPEGAELKPWYEPIYLGGVFQLEKGDQLSAEVNLPKYLDFAE 221
              : :: :|:: :: :: : : | ::|||||||||||||: :||| :| : : :: |
Qy   245 SIKIPSSHNLMKGGSTKNWSGNSEFH-FY-SINVGGFFKLRAGEEISIQVSNPSLLD-PD 301

Db   222 SGQYYFGVI 230
              | ||| |
Qy   302 QDATYFGAF 310

RESULT 7
ID   O35734
AC   O35734;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   TUMOR NECROSIS FACTOR ALPHA PRECURSOR.
GN   TNF-ALPHA.
OS   MARMOTA MONAX (WOODCHUCK).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PBMC;
RA   LOHRENGEL B., LU M., ROGGENDORF M.;
RL   IMMUNOGENETICS 47:332-335(1998).
DR   EMBL; Y14137; E348344; -.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PFAM; PF00229; TNF.
KW   SIGNAL.
FT   SIGNAL   1   77   POTENTIAL.
FT   CHAIN   78   233   POTENTIAL.
SQ   SEQUENCE   233 AA; 25764 MW; 6FCOF34A CRC32;

Query Match   7.0%; Score 161; DB 11; Length 233;
Best Local Similarity   26.2%; Pred. No. 2.76e-09;
  Matches   37; Conservative   34; Mismatches   63; Indels   7; Gaps   7;

Db   94 VAKNEDKEQLV-WLSRRANALLANGMELIDNQLVVPANGLYLVYSQVLFKGQGCPSYVLL 152
              :: ::: : :| | :: :: ||| : : :|||| |::|| : | ||:|| | ||| ||| |
Qy   174 IPSGSHKVTLSSWYHDRGWAKISN-MTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPT 232

Db   153 TH-TVSRFAVSIYQDKV-NLLSAIKSPCPKESLEGAEFKPWYEPIYLGGVFELQKGDRLSA 210
              || || ||: :: ::|: :|:: :: ::: : | :|||||||||||| : ::|| :|
Qy   233 DYLQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFH-FY-SINVGGFFKIRAGEEISI 290

```
Qy   207  VNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYVVKTSIKIPSSHNLMKGGSTK-NWSG 265

Db   192  SSSRVWWDSFLGGVVHLEAGEEVVRVLDERLVRLRDGTRSYFGAFMV 240
          :|: |              |  ||||     ::|: : ::|||||:|
Qy   266  NSEFHFYSIN-VGGFFKLRAGEEISIQVSNPSLLD-PDQDATYFGAFKV 312


RESULT 10
ID   Q28320      PRELIMINARY;     PRT;   234 AA.
AC   Q28320;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   TNF-ALPHA.
OS   CAPRA HIRCUS (GOAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; ARTIODACTYLA.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   TAKAKURA H., MORI Y., TATSUMI M.;
RL   SUBMITTED (JUL-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; D86587; G1483165; -.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PFAM; PF00229; TNF; 1.
SQ   SEQUENCE   234 AA;   25519 MW;   C6424744 CRC32;

Query Match             6.2%;  Score 142;  DB 6;  Length 234;
Best Local Similarity   23.8%;  Pred. No. 2.35e-06;
Matches   43;  Conservative   48;  Mismatches   80;  Indels   10;  Gaps   10;

Db    55  VIGPQREEQSPAGPSFNRPLVQTLRSSSQASSNKPVAHVVANISAPGQ-LRWGDSYAN-A 112
          :||||   |  |   |::   :||  : :::  ||:::||:| ::||  : | |  | :
Qy   133  IVGPQRFSGAPAMMEGSW-LDVAQRGKPEAQPFAHLTINAASIPSGSHKVTLSSWYHDRG 191

Db   113  -LKANGVELKDNQLVVPTDGLYLIYSQVLFRGHGCP-STPL-FLTHTISRIAVSYQTKVN 169
           :| || | ::::  ::  |::||||:|||| ||  :|       :|:|:: |   |:
Qy   192  WAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYVVKTSIKIPSS 251

Db   170  ILSAIKSPCHRETPEGAEAKPWYEPIYQGGVFQLEKGDRLSAEINQPEYLDYAESGQVYF 229
            || :            | ||:    |||||:||| |   ::| :  |  : || || 
Qy   252  -HNLMKGGSTKNWSGNSEFH-FY-SINVGGFFKLRAGEEISIQVSNPSLLD-PDQDATYF 307

Db   230  G 230
          |
Qy   308  G 308


RESULT 11
ID   O70332      PRELIMINARY;     PRT;   216 AA.
AC   O70332;
DT   01-AUG-1998 (TREMBLREL. 07, CREATED)
DT   01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   TUMOR NECROSIS FACTOR-ALPHA (FRAGMENT).
OS   MESOCRICETUS AURATUS (GOLDEN HAMSTER).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SPLEEN;
RA   MELBY P.C., TRYON V.V., CHANDRASEKAR B., FREEMAN G.L.;
RL   INFECT. IMMUN. 66:2135-2142(1998).
DR   EMBL; AF046215; G3005109; -.
DR   PROSITE; PS00251; TNF_1; 1.
FT   NON_TER       1      1
FT   NON_TER     216    216
SQ   SEQUENCE   216 AA;   23793 MW;   79A08367 CRC32;

Query Match             5.4%;  Score 125;  DB 11;  Length 216;
Best Local Similarity   29.2%;  Pred. No. 7.11e-04;
Matches   35;  Conservative   28;  Mismatches   46;  Indels   11;  Gaps    7;

Db    97  WLSHRANALLANGMSLKDNQLVIPADGLYLVYSQVLFRGQ---GC-PS-YVLLTHTVSRI 151
```

```
RESULT 8
ID   Q62326      PRELIMINARY;     PRT;   156 AA.
AC   Q62326;
DT   01-NOV-1996 (TREMBLREL. 01, CREATED)
DT   01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT   01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE   TUMOR NECROSIS FACTOR.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86149365.
RA   CAPUT D., BEUTLER B., HARTOG K., THAYER R., BROWN-SHIMER S.,
RA   CERAMI A.;
RL   PROC. NATL. ACAD. SCI. U.S.A. 83:1670-1674(1986).
DR   EMBL; M13049; E7693; -.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PFAM; PF00229; TNF; 1.
SQ   SEQUENCE   156 AA;   17357 MW;   03975542 CRC32;

Query Match             6.6%;  Score 152;  DB 11;  Length 156;
Best Local Similarity   25.4%;  Pred. No. 7.03e-08;
Matches   32;  Conservative   33;  Mismatches   55;  Indels    6;  Gaps    6;

Db    28  WLSQRANALLANGMDLKDNQLVVPADGLYLVYSQVLFKGQGCPDYVLLTH-TVSRFAISY 86
          ::||||  ||    :  ::: |   :::||  ||:|||||| ||| ||:    :| ::|
Qy   186  WYHDRGWAKISN-MTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYVVKT 244

Db    87  QEKV-NLSAVKSPCPKDTPEGAELKPWYEPIYLGGVFQLEKGDQLSAEVNLPKYLDFAE 145
          :| :  :: |: :||||| :| ||||| ||| ::|:| ||:|:  |  ::|||: |
Qy   245  SIKIPSSHNLMKGGSTKNWSGNSEFH-FY-SINVGGFFKLRAGEEISIQVSNPSLLD-PD 301

Db   146  SGQVYF 151
          || ||
Qy   302  QDATYF 307


RESULT 9
ID   O43557      PRELIMINARY;     PRT;   240 AA.
AC   O43557;
DT   01-JUN-1998 (TREMBLREL. 06, CREATED)
DT   01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT   01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE   TUMOR NECROSIS FACTOR SUPERFAMILY MEMBER LIGHT.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   MAURI D.N., EBNER R., MONTGOMERY R.I., KOCHEL K.D., CHEUNG T.C.,
RA   YU G.-L., RUBEN S., MURPHY M., EISENBERG R.J., COHEN G.H., SPEAR P.G.,
RA   WARE C.F.;
RL   IMMUNITY 8:21-30(1998).
DR   EMBL; AF036581; G2815624; -.
SQ   SEQUENCE   240 AA;   26351 MW;   4A4B603A CRC32;

Query Match             6.5%;  Score 148;  DB 4;  Length 240;
Best Local Similarity   27.2%;  Pred. No. 2.89e-07;
Matches   42;  Conservative   46;  Mismatches   71;  Indels   10;  Gaps   10;

Db    78  GSWEQLIQERRSHEVNPAAHLTCANSSL-TGS-GGPLL-WETQLGLAFLRGLSYHDGALV 134
          |||| ||     ||    |||| |    ||:  ::|  :|| :|||||| |  ::| |
Qy   148  GSWLDVAQ-RGKPEAQPFAHLTINAASIPSGSHKVTLSSWYHDRGWAKISNMTLSNGKLR 206

Db   135  VTKAGYYYIYSKVQLGGVGCPLGLASTITHGL-YK-RTP-RYPEELELLVSQQSPCGRAT 191
          | ||:||||::      :::   :  ::: | |  :|   |:::   :|  :  : ::
```

```
Qy   87 YQTKV-NLSAIKSPCQRETPRGAEAKPWYEPIYILGGVFQLEKGDRLSAEINRPDYLDFA 145
         |: :|: :|  ::|: : :: || :  |: || | :|::|| ||::|:  ::|
Db  244 TSIKIPSSHNLMKGGSTKNWSGNSEFH-FY-SINVGGFFKLRAGEEISIQVSNPSLLD-P 300

Db  146 ESGQVYFG 153
         :  ||
Qy  301 DQDATYFG 308

RESULT 14
ID   O54967              PRELIMINARY;          PRT;   1055 AA.
AC   O54967;
DT   01-JUN-1998 (TREMBLREL. 06, CREATED)
DT   01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   NON-RECEPTOR PROTEIN TYROSINE KINASE ACK.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; RODENTIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RA   HER J.-H., BOLEN J.B.;
RL   SUBMITTED (DEC-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF037260; G2921447; -.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
SQ   SEQUENCE  1055 AA;  116834 MW;  F1AB2646 CRC32;

Query Match             5.2%;   Score 120;  DB 11;  Length 1055;
Best Local Similarity  40.5%;   Pred. No. 3.56e-03;
 Matches  17;  Conservative  12;  Mismatches  12;  Indels  1;  Gaps  1;

Db  887 YQRFLREAQSPEEPAALP-VPPLLPPPSTPAPAAPTATVRPM 927
         | : :|| :  ::  |:  || : :|||||||||
Qy    8 YGKYLRSSEEMGSGPGVPHEGPLHPAPSAPAPPPAASRSM 49

RESULT 15
ID   O02755              PRELIMINARY;          PRT;    348 AA.
AC   O02755;
DT   01-JUL-1997 (TREMBLREL. 04, CREATED)
DT   01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE   CCAAT/ENHANCER-BINDING BETA PROTEIN.
GN   C/EBP BETA.
OS   BOS TAURUS (BOVINE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; ARTIODACTYLA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAPANESE BLACK CATTLE;
RX   MEDLINE; 97203913.
RA   YAMAOKA I., TANIGUCHI Y., SASAKI Y.;
RL   J. ANIM. SCI. 75:587-587(1997).
DR   EMBL; D82985; D1020911; -.
SQ   SEQUENCE  348 AA;  36390 MW;  76720BD9 CRC32;

Query Match             4.6%;   Score 106;  DB 6;  Length 348;
Best Local Similarity  23.3%;   Pred. No. 2.64e-01;
 Matches  28;  Conservative  35;  Mismatches  49;  Indels  8;  Gaps  8;

Db   74 DFSPYLEPLGAPQAPAPTASD-TFEAAPSAPAVPASSGQHHDFLSDLFSDDYGGKNCK 132
         |: :|| : :::| ::||||||  :|||||| || :
Qy    7 DYGKYLRS-SEEMGSGPGVPHEGPLHPAPSAPAPAPPPAASRSMFLA-LLGLGLGQVVCS 64

Db  133 KAAEYGYVSLGRLGAAKGALHPG-CFAP-LHPPPPPPPPAELKAEPGFEPADCKRKEEA 190
         |: || :: :: | |:| | : ||::|  | ||::|!
Qy   65 IAL-FLYFR-AQMDPNRISEDSTHCFYRILRLHENAGLQDSTLESEDTL-PDSCRRMKQA 121

Search completed: Fri Nov 13 17:18:57 1998
```

```
Qy  186 WYHDRGWAKISN-MTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYVVKT 244
         | :  |  | :  |:  ::| ||  ::|  |: || | :  ::|: | :|  | :|
Db  152 AVSYEDNVNLLSAIKSPCPKETPEGEELKPWYEPIYLGGVFQLEKGDRLSAEVNLPKYLD 211
         ::  :| :  |:  |:  ||   :|  || :
Qy  245 SIKIPSSHNLM--KGGSTKNWSGNSEFH-FY-SINVGGFFKLRAGEEISIQVSNPSLLD 299

RESULT 12
ID   O18779              PRELIMINARY;          PRT;    233 AA.
AC   O18779;
DT   01-JAN-1998 (TREMBLREL. 05, CREATED)
DT   01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT   01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE   TUMOR NECROSIS FACTOR ALPHA.
GN   TNFA.
OS   BOS TAURUS (BOVINE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; ARTIODACTYLA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N'DAMA;
RA   IRAQI F.;
RL   SUBMITTED (JUL-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF011926; G2599343; -.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PFAM; PF00229; TNF.
SQ   SEQUENCE  233 AA;  25395 MW;  64FAFC88 CRC32;

Query Match             5.4%;   Score 123;  DB 6;  Length 233;
Best Local Similarity  25.2%;   Pred. No. 1.36e-03;
 Matches  32;  Conservative  29;  Mismatches  61;  Indels  5;  Gaps  5;

Db  104 WWDSYANALMANGVKLEDNQLVVPADGLYLIYSQVLFRGQGCPSTPLFTHTISRIAVSY 163
         | :  :  ::| |  || ::| : :| ||:  |: || |  :
Qy  186 WYHDRGWAKISN-MTLSNGKLRVNQDGFYYLYANICFRHHE-TSGSV-PTDYLQLMVYVV 242

Db  164 QTKVNILSAIKSPCHRETPEWAEAKPW-YEPIYQGVFQLEKGDRLSAEINLPDYLDYAE 222
         :|: :|::|| | :|   |   ||   || | ||||:|| | |:::| || | |
Qy  243 KTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLD-PD 301

Db  223 SGQVYFG 229
         | ||
Qy  302 QDATYFG 308

RESULT 13
ID   O43647              PRELIMINARY;          PRT;    157 AA.
AC   O43647;
DT   01-JUN-1998 (TREMBLREL. 06, CREATED)
DT   01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT   01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE   TUMOR NECROSIS FACTOR ALPHA (FRAGMENT).
GN   TNFA.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC   EUTHERIA; PRIMATES.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   JANG J.S., KIM B.E.;
RL   SUBMITTED (JAN-1998) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR   EMBL; AF043342; G2905634; -.
DR   PROSITE; PS00251; TNF_1; 1.
FT   NON_TER       1        1
SQ   SEQUENCE  157 AA;  17380 MW;  FD78F0A6 CRC32;

Query Match             5.3%;   Score 122;  DB 4;  Length 157;
Best Local Similarity  24.2%;   Pred. No. 1.88e-03;
 Matches  31;  Conservative  32;  Mismatches  58;  Indels  7;  Gaps  7;

Db   28 WLNRRANALLANGVELRDNQLVPSEGLLIYSQVLFKGQGCPSTHVLLTHTISRI-AVS 86
         | :  : | ||| ||| :||| |  |: || |  || |  :|
Qy  186 WYHDRGWAKISN-MTLSNGKLRVNQDGFYYLYANICFRHHE-TSGSVPTDYLQLMVYVVK 243
```

US-08-989-362-2.rspt

Job time : 131 secs.

```
*************************************************************
*                                                           *
*   |\/| |_) (_  |_) (_ |_|                                 *
*   |  | |   __) | \ __) | |  (TM)                          *
*                                                           *
*************************************************************

        Release 3.1A John F. Collins, Biocomputing Research Unit.
        Copyright (c) 1993-1998 University of Edinburgh, U.K.
                 Distribution rights by Oxford Molecular Ltd

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:          Fri Nov 13 17:10:41 1998;  MasPar time 9.91 Seconds
                                  515.747 Million cell updates/sec
Tabular output not generated.

Title:          >US-08-989-362-2
Description:    (1-316) from US08989362.pep
Perfect Score:  2294
Sequence:       1 MRRASRDYGKYLRSSEEMGS.........LLDPDQDATYFGAFKVQDID 316

Scoring table:  PAM 150
                Gap 11

Searched:       131922 seqs, 16180660 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:       a-geneseq32
                1:part1  2:part2  3:part3  4:part4  5:part5  6:part6  7:part7
                8:part8  9:part9 10:part10 11:part11 12:part12 13:part13
               14:part14 15:part15 16:part16 17:part17 18:part18
               19:part19 20:part20 21:part21 22:part22 23:part23
               24:part24 25:part25 26:part26 27:part27 28:part28
               29:part29

Statistics:     Mean 33.982;  Variance 147.062;  scale 0.231

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

                              SUMMARIES

Result          Query
No.   Score  Match Length DB  ID      Description              Pred. No.
----  -----  ----- ------ --- ------  -----------              --------
 1     326   14.2   281   27  W27134  Human Apoptosis induc    5.41e-19
 2     326   14.2   281   29  W44354  Human AGP-1.             5.41e-19
 3     326   14.2   281   23  W19777  Novel cytokine Apo-2     5.41e-19
 4     326   14.2   281   23  W19787  Human apoptosis induc    5.41e-19
 5     282   12.3   291   29  W44353  Murine AGP-1.            5.30e-15
 6     282   12.3   291   23  W19788  Mouse apoptosis induc    5.30e-15
 7     176    7.7   182   23  W09128  CD40 ligand/zipper do    9.75e-06
 8     174    7.6   137   14  R79065  Rat Fas ligand (parti    1.43e-05
 9     174    7.6   138   15  R79063  Rat Fas ligand (parti    1.43e-05
10     174    7.6   178   14  R88308  C-terminally deleted     1.43e-05
11     175    7.6   179   23  W11814  Fas ligand.              1.18e-05
12     175    7.6   179   14  R79099  Human Fas ligand (par    1.18e-05
13     174    7.6   179   14  R79066  Rat Fas ligand (parti    1.43e-05
14     174    7.6   278   14  R79095  Rat Fas ligand encode    1.43e-05
15     175    7.6   281   26  W27143  Human Fas ligand.        1.18e-05
16     175    7.6   281   14  R77281  Human Fas-L protein.     1.18e-05
17     175    7.6   281   14  R79097  Human Fas ligand.        1.18e-05
18     175    7.6   281   19  R98104  Human Fas ligand deri    1.18e-05
```

```
19  281  18  R88356  Human Fas ligand.              1.43e-05
20  287  26  W27146  CH3/Fas ligand domain          1.18e-05
21  137  14  R79067  Mouse Fas ligand (par          2.11e-05
22  138  14  R79068  Mouse Fas ligand (par          2.11e-05
23  179  14  R79069  Mouse Fas ligand (par          2.11e-05
24  279  14  R79098  Mouse Fas ligand.              2.11e-05
25  279  18  R88357  Mouse Fas ligand.              3.10e-05
26  279  14  R77282  Mouse Fas-L protein.           3.10e-05
27  169  23  W09120  CD40 ligand/zipper do          3.75e-05
28  130  15  R79096  Human Fas ligand C-te          4.54e-05
29  136  14  R88307  N-terminally deleted           4.54e-05
30  168  14  R88306  N-terminally deleted           4.54e-05
31  168  14  R79100  Human Fas ligand (par          4.54e-05
32  168  14  R88305  N-terminally deleted           4.54e-05
33  168  15  R79064  Human Fas ligand (par          4.54e-05
34  139  14  R88304  N-terminally deleted           4.54e-05
35  141  14  R88303  N-terminally deleted           4.54e-05
36  168  23  W16667  Human Fas ligand (hFa          4.54e-05
37  167  23  W10875  FasL/LT-alpha hybrid.          5.50e-05
38  166  23  W24009  CD40 ligand/zipper do          6.66e-05
39  166  23  W24013  CD40 ligand/zipper do          6.66e-05
40  166   9  R49548  Predicted sequence en          6.66e-05
41  166   7  W09113  Human CD40L mutein C1          6.66e-05
42  165  11  R57469  CD40 ligand.                   8.06e-05
43  165  10  R53969  Human CD40-L type II           8.06e-05
44  166  23  W09129  CD40 ligand/zipper do          6.66e-05
45  378  27  W35864  Human FAS-ligand:IgG2          6.66e-05
```

```
                          ALIGNMENTS

RESULT 1
ID   W27134 standard; Protein; 281 AA.
AC   W27134;
DT   02-APR-1998  (first entry)
DE   Human Apoptosis inducing molecule-I (AIM-I).
KW   Apoptosis inducing molecule-I; AIM-I; autoimmune disorder;
KW   tumour necrosis factor ligand superfamily; AIM-I altered expression;
KW   neoplasia inhibition; anti-inflammatory agent.
OS   Homo sapiens.
PN   WO973899-A1.
PD   18-SEP-1997.
PF   14-MAR-1996; U03773.
PR   14-MAR-1996; WO-U03773.
PA   (HUMA-) HUMAN GENOME SCI INC.
PI   Ruben SM;
DR   WPI; 97-470807/43.
DR   N-PSDB; T85210.
PT   New isolated apoptosis inducing molecule-I - used to develop
PT   products for the diagnosis and therapy of e.g. autoimmune diseases,
PT   tumours; graft versus host disease or inflammation.
PS   Claim 2; Fig 1; 82pp; English.
CC   The present sequence represents a human Apoptosis inducing molecule-I
CC   (AIM-I). AIM-I is a member of the tumour necrosis factor (TNF) ligand
CC   superfamily. The products can be used in the diagnosis and treatment of
CC   disorders related to under-expression, over-expression or altered
CC   expression of AIM-I. AIM-I or agonists can be used for treating
CC   autoimmune disorders including systemic lupus erythematosus,
CC   immunoproliferative disease lymphadenopathy (IPL),
CC   angioimmunoproliferative lymphadenopathy (AIL), rheumatoid arthritis,
CC   diabetes, and multiple sclerosis, graft versus host disease, to inhibit
CC   neoplasia such as tumour cell growth, to treat restenosis, to regulate
CC   haematopoiesis in endothelial cell development, to stimulate peripheral
CC   tolerance and cytotoxic T-cell mediated apoptosis. Antagonists can be
CC   used for treating cachexia, cerebral malaria, rheumatoid arthritis
CC   or osteoporosis, for preventing graft-host rejection, and as
CC   anti-inflammatory agents, for treating endotoxic shock or to prevent
CC   activation of HIV.
SQ   Sequence 281 AA;

Query Match        14.2%;       Score 326;   DB 27;   Length 281;
Best Local Similarity   35.9%;       Pred. No. 5.41e-19;
Matches   47;  Conservative  33;  Mismatches  49;  Indels  2;  Gaps  2;
```

```
Db   151 inswessrsghsflsnlhlrngelvihekgfyyiysqtyfrfgeeikentkndkqmvqyl 210
         ::||:||-|-|-::||:||-:|||||:-||||:-||:
Qy   183 LSSWYHDR-GWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYV 241

Db   211 ykyt-sypdpillmksarnscwskdaeyglysiyqggifelkendrifvsvtnehlidmd 269
         |-:||:-|||:|-||-:-|||-:||-:|:|||:|-:::||-|||
Qy   242 VKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDPD 301

Db   270 heasffgaflv 280
         ::|:|||||
Qy   302 QDATYFGAFKV 312

RESULT   2
ID   W44354 standard; Protein; 281 AA.
AC   W44354;
DT   28-MAY-1998  (first entry)
DE   Human AGP-1.
KW   Human; AGP-1; tumour necrosis factor-related protein; TNF;
KW   inflammation; bone resorption; haematopoietic disease.
OS   Homo sapiens.
PN   WO9746686-A2.
PD   11-DEC-1997.
PF   06-JUN-1997; U09895.
PR   07-JUN-1996; US-660562.
PA   (AMGE ) AMGEN INC.
PI   Danilenko DM, Johnson MJ, Simonet WS;
DR   WPI; 98-042194/04.
DR   N-PSDB; V15295.
PT   Nucleic acid encoding AGP-1, a tumour necrosis factor-related
PT   protein - useful for treating inflammation, bone resorption and
PT   haematopoietic diseases
PS   Claim 7; Page 36-37; 54pp; English.
CC   The present sequence represents human AGP-1. AGP-1 is a tumour-necrosis
CC   factor (TNF)-related protein, involved in inflammation, myelopoiesis
CC   and bone resorption. It has the same nucleic acid and amino acid (aa)
CC   sequences as the TNF-related apoptosis-induced ligand (TRAIL) described
CC   in Immunity, 3 (1995) 673. Antibodies (Ab) are used as immunoassay
CC   reagents for detecting AGP-1 expression. Nucleic acid complementary to
CC   AGP-1 is used to regulate AGP-1 expression and antagonistic compounds
CC   are used to treat inflammation (e.g. rheumatoid arthritis, systemic
CC   lupus erythematosus, psoriasis, scleroderma, infection-related
CC   inflammation) or bone resorption diseases (e.g. osteoporosis,
CC   osteomyelitis, hypercalcaemia, Paget's disease). AGP-1 can be used to
CC   treat haematopoietic diseases associated with reduction in the number
CC   of bone marrow cells, particularly neutrophils and lymphocytes, e.g.
CC   where caused by disease, injury or exposure to myelosuppressive agents.
CC   Host cells, transformed with expression vectors containing AGP-1 DNA,
CC   are used to produce recombinant AGP-1.
SQ   Sequence 281 AA;

Query Match             14.2%;  Score 326; DB 29; Length 281;
Best Local Similarity   35.9%;  Pred. No. 5.41e-19;
Matches    47;  Conservative   33;  Mismatches  49;  Indels    2;  Gaps   2;

Db   151 inswessrsghsflsnlhlrngelvihekgfyyiysqtyfrfgeeikentkndkqmvqyl 210
         ::||:||-|-|-::||:||-:|||||:-||||:-||:
Qy   183 LSSWYHDR-GWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYV 241

Db   211 ykyt-sypdpillmksarnscwskdaeyglysiyqggifelkendrifvsvtnehlidmd 269
         |-:||:-|||:|-||-:-|||-:||-:|:|||:|-:::||-|||
Qy   242 VKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDPD 301

Db   270 heasffgaflv 280
         ::|:|||||
Qy   302 QDATYFGAFKV 312

RESULT   3
ID   W19777 standard; Protein; 281 AA.
AC   W19777;
DT   22-SEP-1997  (first entry)
DE   Novel cytokine Apo-2 ligand.
KW   Apo-2 ligand; cytokine; apoptosis; breast cancer; colon cancer;
KW   therapy.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   peptide         1..281
FT                   /note= "Claim 4"
FT   protein         15..281
FT                   /note= "Claim 3"
FT   protein         41..281
FT                   /note= "Claim 2"
FT   protein         114..281
FT                   /note= "Claim 1"
FT   region          1..14
FT                   /label= Cytoplasmic_region
FT   region          15..40
FT                   /label= Transmembrane_region
FT   region          41..281
FT                   /label= Extracellular_region
FT   modified_site   109
FT                   /label= Glycosylation
FT                   /note= "putative N-linked glycosylation site"
PN   WO9725428-A1.
PD   17-JUL-1997.
PF   08-JAN-1997; U00272.
PR   09-JAN-1996; US-584031.
PA   (GETH ) GENENTECH INC.
PI   Ashkenazi AJ, Chuntharapai A,  Kim KJ;
DR   WPI; 97-372867/34.
DR   N-PSDB; T72796.
PT   Novel cytokine, Apo-2 ligand and corresponding DNA - used to induce
PT   apoptosis for the treatment of breast and colon cancer
PS   Claim 4; Fig 1a; 72pp; English.
CC   A novel cytokine (W19777), designated Apo-2 ligand, induces
CC   mammalian cell apoptosis.  It is believed to be a member of the
CC   tumour necrosis factor cytokine family.  Its amino acid sequence
CC   was deduced from a cDNA clone (T72796) isolated from a human
CC   placental cDNA library.  Apo-2 ligand polypeptides, esp. those
CC   corresponding to amino acids 114-281, 41-281, 15-281 or 1-281 of
CC   Apo-2 ligand, can be produced by culturing cells transformed or
CC   transfected with a vector contg. Apo-2 ligand nucleic acid.  They
CC   can be used to induce apoptosis in mammals and to treat
CC   pathological conditions such as cancer (esp. breast or colon
CC   cancer) or to raise antibodies useful in diagnostic assays.
SQ   Sequence 281 AA;

Query Match             14.2%;  Score 326; DB 23;  Length 281;
Best Local Similarity   35.9%;  Pred. No. 5.41e-19;
Matches    47;  Conservative   33;  Mismatches  49;  Indels    2;  Gaps   2;

Db   151 inswessrsghsflsnlhlrngelvihekgfyyiysqtyfrfgeeikentkndkqmvqyl 210
         :||:||-|-|-::||:||-:|||||:-||||:-||:
Qy   183 LSSWYHDR-GWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYV 241

Db   211 ykyt-sypdpillmksarnscwskdaeyglysiyqggifelkendrifvsvtnehlidmd 269
         |-:||:-|||:|-||-:-|||-:||-:|:|||:|-:::||-|||
Qy   242 VKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDPD 301

Db   270 heasffgaflv 280
         ::|:|||||
Qy   302 QDATYFGAFKV 312

RESULT   4
ID   W19787 standard; Protein; 281 AA.
AC   W19787;
DT   24-SEP-1997  (first entry)
DE   Human apoptosis inducer cytokine TRAIL.
KW   Tumour necrosis factor related apoptosis inducing ligand; TRAIL;
KW   cytokine; cancer; leukaemia; lymphoma; melanoma; viral infection;
KW   thrombotic microangioplasty; therapy.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
```

```
FT   domain          1..18
FT                   /label= Cytoplasmic_domain
FT   domain          19..38
FT                   /label= Transmembrane_domain
FT   domain          39..281
FT                   /label= Extracellular_domain
FT                   /note="contains a receptor-binding region"
FT   modified_site   109..111
FT                   /note="potential N-glycosylation site"
FT   cleavage_site   89..90
FT                   /note="potential KEX2 protease processing site"
FT   cleavage_site   149..150
FT                   /note="potential KEX2 protease processing site"
PN   WO9701633-A1.
PD   16-JAN-1997.
PF   25-JUN-1996; U10895.
PR   29-JUN-1995; US-496632.
PR   01-NOV-1995; US-548368.
PA   (IMMV ) IMMUNEX CORP.
PI   Goodwin RG, Wiley SR;
DR   WPI: 97-118715/11.
DR   N-PSDB; T72847.
PT   TRAIL, a novel cytokine, induces apoptosis in cancer and
PT   virus-infected cells - useful for treating thrombotic
PT   microangiopathy, cancer and viral infection and for use in assays
PS   Claim 10; Page 43-44; 62pp; English.
CC   Human tumour necrosis factor related apoptosis inducing ligand
CC   (TRAIL) (W19787) is a novel cytokine that induces apoptosis of
CC   certain target cells, including cancer cells and virally infected
CC   cells. Its amino acid sequence was deduced from cDNA clone HuAIC
CC   (T72848), deposited in vector pDC409 as ATCC 69849.  Recombinant
CC   TRAIL polypeptides (esp. soluble polypeptides) can be expressed
CC   in host cells and used in the treatment of cancer (e.g. leukaemia,
CC   lymphoma and melanoma) and viral infections, or to raise antibodies
CC   that may be useful for treating thrombotic microangiopathies.
SQ   Sequence 281 AA;

Query Match             14.2%;  Score 326;  DB 23;  Length 281;
Best Local Similarity   35.9%;  Pred. No. 5.41e-19;
  Matches   47;  Conservative   33;  Mismatches  49;  Indels  2;  Gaps  2;

Db  151 inswessrsghsflsnlhlrngelvihekgfyyiysqtyfrfgeeikentkndqmvqyi 210
            ::||:|:|:::||:|||:||::|||||:|| :|  :|::|
Qy  183 LSSWYHDR-GWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMYV 241

Db  211 ykyt-sypdpillmksarnscwskdaeyglysiyqggifelkendrifvsvtnehlidmd 269
            ||:| |||:|| ||:|| ::|:||| :||:|:|::|:| | :|:| |||:| :|:|
Qy  242 VKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDPD 301

Db  270 heasffgaflv 280
            |:||||||:|
Qy  302 QDATYFGAFKV 312

RESULT 5
ID   W44353 standard; Protein; 291 AA.
AC   W44353;
DT   28-MAY-1998 (first entry)
DE   Murine AGP-1.
KW   Murine; AGP-1; tumour necrosis factor-related protein; TNF;
KW   inflammation; bone resorption; haematopoietic disease.
OS   Mus sp.
PN   WO9746686-A2.
PD   11-DEC-1997.
PF   06-JUN-1997; U09895.
PR   07-JUN-1996; US-660562.
PA   (AMGE-) AMGEN INC.
PI   Danilenko DM, Johnson MJ, Simonet WS;
DR   WPI: 98-042194/04.
DR   N-PSDB; V15294.
PT   Nucleic acid encoding AGP-1, a tumour necrosis factor-related
PT   protein - useful for treating inflammation, bone resorption and
PT   haematopoietic diseases
```

```
PS   Claim 7; Page 33-34; 54pp; English.
CC   The present sequence represents murine AGP-1. AGP-1 is a tumour-necrosis
CC   factor (TNF)-related protein, involved in inflammation, myelopoiesis
CC   and bone resorption. It has the same nucleic acid and amino acid (aa)
CC   sequences as the TNF-related apoptosis-induced ligand (TRAIL) described
CC   in Immunity, 3 (1995) 673. Antibodies (Ab) are used as immunoassay
CC   reagents for detecting AGP-1 expression. Nucleic acid complementary to
CC   AGP-1 is used to regulate AGP-1 expression and antagonistic compounds
CC   are used to treat inflammation (e.g. rheumatoid arthritis, systemic
CC   lupus erythematosus, psoriasis, scleroderma, infection-related
CC   inflammation) or bone resorption diseases (e.g. osteoporosis,
CC   osteomyelitis, hypercalcaemia, Paget's disease). AGP-1 can be used to
CC   treat haematopoietic diseases associated with reduction in the number
CC   of bone marrow cells, particularly neutrophils and lymphocytes, e.g.
CC   where caused by disease, injury or exposure to myelosuppressive agents.
CC   Host cells, transformed with expression vectors containing AGP-1 DNA,
CC   are used to produce recombinant AGP-1.
SQ   Sequence 291 AA;

Query Match             12.3%;  Score 282;  DB 29;  Length 291;
Best Local Similarity   35.6%;  Pred. No. 5.30e-15;
  Matches   48;  Conservative   37;  Mismatches  42;  Indels  8;  Gaps  5;

Db  155 ieswessrkghsflnhvlfrngelviegeglyyiysqtyfrfgeaedaskmvskdkvrtk 214
            :| |:|  ::::|:|||:|||::|:::|:||||:|:||||:| :|  :|:::|
Qy  183 LSSWYHDR-GWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHET-SGS--VPTDYL--- 235

Db  215 qlvqviykyt-sypdpivlmksarnscwsrdaeyglysiyqgglfelkkndrifvsvtne 273
            ||  : ||:| |||:|| ||:||::|:  :|:||| :||:|::|::|: :|:| | :|:|
Qy  236 QLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNP 295

Db  274 hlmdldqeasffgaf 288
            |:||:||||||||:|
Qy  296 SLLDPDQDATYFGAF 310

RESULT 6
ID   W19788 standard; Protein; 291 AA.
AC   W19788;
DT   24-SEP-1997 (first entry)
DE   Mouse apoptosis inducer cytokine TRAIL.
KW   Tumour necrosis factor related apoptosis inducing ligand; TRAIL;
KW   cytokine; cancer; leukaemia; lyphoma; melanoma; viral infection;
KW   thrombotic microangioplasty; therapy.
OS   Mus musculus.
FH   Key              Location/Qualifiers
FT   domain           1..17
FT                    /label= Cytoplasmic_domain
FT   domain           18..38
FT                    /label= Transmembrane_domain
FT   domain           39..291
FT                    /label= Extracellular_domain
FT   modified_site    52..54
FT                    /note="potential N-glycosylation site"
FT   cleavage_site    85..86
FT                    /note="potential KEX2 protease processing site"
FT   cleavage_site    135..136
FT                    /note="potential KEX2 protease processing site"
FT   cleavage_site    162..163
FT                    /note="potential KEX2 protease processing site"
PN   WO9701633-A1.
PD   16-JAN-1997.
PF   25-JUN-1996; U10895.
PR   29-JUN-1995; US-496632.
PR   01-NOV-1995; US-548368.
PA   (IMMV ) IMMUNEX CORP.
PI   Goodwin RG, Wiley SR;
DR   WPI: 97-118715/11.
DR   N-PSDB; T72848.
PT   TRAIL, a novel cytokine, induces apoptosis in cancer and
PT   virus-infected cells - useful for treating thrombotic
PT   microangiopathy, cancer and viral infection and for use in assays
PS   Claim 10; Page 48-49; 62pp; English.
```

CC   Mouse tumour necrosis factor related apoptosis inducing ligand
CC   (TRAIL) (W19788) is a novel cytokine that induces apoptosis of
CC   certain target cells, including cancer cells and virally infected
CC   cells. Its amino acid sequence was deduced from cDNA clone MuAIC
CC   (T72849), isolated from a T cell line 7B9 library. Human TRAIL
CC   (W19787) has also been identified. Recombinant TRAIL polypeptides
CC   (esp. soluble polypeptides) can be expressed in host cells and used
CC   in the treatment of cancer (e.g. leukaemia, lymphoma and melanoma)
CC   and viral infections, or to raise antibodies that may be useful for
CC   treating thrombotic microangiopathies.
SQ   Sequence 291 AA;

Query Match   12.3%; Score 282; DB 23; Length 291;
Best Local Similarity   35.6%; Pred. No. 5.30e-15;
Matches   48; Conservative   37; Mismatches   42; Indels   8; Gaps   5;

Db   155 leswessrkghsflnhvlfrngelviegeglyyiysqtyfrfqeeedaskmvskdvrtk 214
                : || | ||.| ::::| ||:|:||::|:. :| :|:| :
Qy   183 LSSWYHDR-GWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHET-SGS--VPTDYL--- 235

Db   215 qlvqiyikyt-sypdpivimksarnscwsrdaeyyglysiyqgglfelkndrifvsvtne 273
                ||:|:-|: || :||:|:| ||:|| |||:|:|
Qy   236 QLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHYSINVGGFFKLRAGEEISIQVSNP 295

Db   274 hlmdldqeasffgaf 288
                :|| |||::::|||
Qy   296 SLLDPDQDATYFGAF 310

RESULT 7
ID   W09128 standard; Protein; 182 AA.
AC   W09128;
DT   24-SEP-1997 (first entry)
DE   CD40 ligand/zipper domain fusion protein mutant 255F.
KW   CD40 ligand; membrane bound glycoprotein; B cell proliferation;
KW   antibody secretion; immunoglobulin E; cytokine; CD40L; mutein;
KW   oligomerisation domain; fusion protein; dimer; trimer.
OS   Chimeric - Homo sapiens.
OS   Chimeric - Saccharomyces cerevisiae.
OS   Synthetic.
FH   Key   Location/Qualifiers
FT   domain   1..33
FT   /label= leucine_zipper
FT   /note= "Yeast GCN4 leucine zipper"
FT   region   34..182
FT   /label= CD40L_extracellular_fragment
FT   /note= "Corresponds to residues 113..261 of CD40L
FT   with a S255F mutation"
FT   misc_difference 176
FT   /note= "Wild-type Ser has been replaced by Phe"

PN   WO940918-A2.
PD   19-DEC-1996.
PF   06-JUN-1996; U09632.
PR   07-JUN-1995; US-484624.
PR   07-JUN-1995; US-477733.
PA   (IMMV) IMMUNEX CORP.
PI   Armitage RJ, Fanslow WC, Gibson MG, Spriggs MK;
PI   Srinivasan S;
WPI; 97-052320/05.
DR   WPI; 97-052320/05.
PT   New CD40 ligand mutein with higher CD40 affinity than native ligand
PT   - useful in binding assays, and for therapy of disorders and
PT   diseases involving low levels of B cells and antibody secretion
PS   Example 4: Page -; Page -; 31pp; English.
CC   This sequence represents a fusion of the yeast GCN4 leucine zipper
CC   domain and amino acids 113-261 of the human CD40 ligand containing
CC   a random mutation. The resulting CD40L mutein did not bind to the
CC   CD40 receptor.
CC   (Note: The present sequence does not appear in the specification;
CC   it has been produced using the sequences of the leucine zipper and
CC   wild-type CD40L which are given on page 23 and pages 20-21,
CC   respectively).
SQ   Sequence 182 AA;

Query Match   7.7%; Score 176; DB 23; Length 182;
Best Local Similarity   29.5%; Pred. No. 9.75e-06;
Matches   44; Conservative   48; Mismatches   43; Indels   14; Gaps   12;

Db   43 laahviseasskttsvlqw-aekgyytmsnnlvtlengkqltvkrgglyyiyaqvtfcsn 101
                | ||:||.:: |::| :|| || ||.| : | -| |
Qy   169 INAASIPSGSHKVT-LSSWYHDRGWAKISN-M-TLSNGK-LRVNQDGFYYLYANICF-RH 223

Db   102 reassqap--fiasl-cl-kspgrf--erillraanths-sakpcgq-qsihlggvfelq 153
                :|:|| :| :|:|: :| |::::: |||| |: ||||:
Qy   224 HETSGSVPTDYIQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHYSINVGGFFKLR 283

Db   154 pgasvfnvtdpsqvshgtgtffgilkl 182
                :| :|:|:: |:||
Qy   284 AGEEISIQVSNPSLLDPDQDATYFGAFKV 312

RESULT 8
ID   R79065 standard; Protein; 137 AA.
AC   R79065;
DT   22-FEB-1996 (first entry)
DE   Rat Fas ligand (partial sequence).
KW   Fas ligand; Tumour Necrosis factor family; apoptosis; cell death;
KW   Fas cell surface antigen; Fas-L; rat.
OS   Rattus rattus.
PN   WO9513293-A1.
PD   18-MAY-1995.
PF   10-NOV-1994; J01899.
PR   10-NOV-1993; JP-305975.
PR   13-DEC-1993; JP-342526.
PR   18-MAR-1994; JP-074344.
PR   08-JUL-1994; JP-180955.
PR   07-SEP-1994; JP-239363.
PR   18-OCT-1994; JP-278378.
PA   (MOCH) MOCHIDA PHARM CO LTD.
PA   (OSAB) OSAKA BIOSCIENCE INST.
PI   Nagata S, Nakamura N, Suda T, Takahashi T;
DR   WPI; 95-194031/25.
DR   N-PSDB; Q99494.
PT   Peptide which binds to Fas antigen, and antibody reactive with it -
PT   for treatment and diagnosis of viral or auto:immune diseases
PS   Claim 5: Page 212-213; 300pp; Japanese.
CC   Fas ligands or active fragments able to induce apoptosis in cells
CC   which express the Fas cell surface antigen are claimed. The
CC   proteins are isolated from human, rat and mouse sources. The present
CC   sequence represents part of the rat Fas ligand.
SQ   Sequence 137 AA;

Query Match   7.6%; Score 174; DB 14; Length 137;
Best Local Similarity   26.7%; Pred. No. 1.43e-05;
Matches   36; Conservative   38; Mismatches   53; Indels   8; Gaps   6;

Db   11 srsiple-wedtygtalisgvkykkgglvineaglyfvyskvyfrgq-scnsqp---sh 65
                || :|:::|||| | ::| |: ::| .:| |
Qy   178 SHKVTLSSWYHDRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSSVPTDYLQL 237

Db   66 kvym-r-nfkypgdvlmee-kklnycttgqiwahssylgavfnltvadhlyvnisqlsl 122
                :||: : || ::| |:| :|:||
Qy   238 MVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHYSINVGGFFKLRAGEEISIQVSNPSL 297

Db   123 lnfeesktffglykl 137
                ::: |:|| :|:
Qy   298 LDPDQDATYFGAFKV 312

RESULT 9
ID   R79063 standard; Protein; 138 AA.
AC   R79063;
DT   21-FEB-1996 (first entry)
DE   Rat Fas ligand (partial sequence).
KW   Fas ligand; Tumour Necrosis factor family; apoptosis; cell death;
OS   Rattus rattus.

```
PN  WO9513293-A1.
PD  18-MAY-1995.
PF  10-NOV-1994; JP-305975.
PR  10-NOV-1993; JP-305975.
PR  13-DEC-1993; JP-342526.
PR  18-MAR-1994; JP-074344.
PR  08-JUL-1994; JP-180955.
PR  07-SEP-1994; JP-239363.
PR  18-OCT-1994; JP-278378.
PA  (MOCH ) MOCHIDA PHARM CO LTD.
PA  (OSAB-) OSAKA BIOSCIENCE INST.
PI  Nagata S, Nakamura N, Suda T, Takahashi T;
DR  WPI; 95-194031/25.
DR  N-PSDB; Q99492.
PT  Peptide which binds to Fas antigen, and antibody reactive with it  -
PT  for treatment and diagnosis of viral or auto:immune diseases
PS  Claim 6; Page 213-215; 300pp; Japanese.
CC  Fas ligands or active fragments able to induce apoptosis in cells
CC  which express the Fas cell surface antigen are claimed. The
CC  proteins are isolated from human, rat and mouse sources. The present
CC  sequence represents part of the rat Fas ligand.
SQ  sequence 138 AA;

Query Match       7.68;  Score 174;  DB 15;  Length 138;
Best Local Similarity 26.78;  Pred. No. 1.43e-05;
Matches  36;  Conservative  38;  Mismatches  53;  Indels  8;  Gaps  6;

Db   12  srsiple-wedtygtalisgvkykgglvineaglyfvyskvyfrgq-scnsqp---lsh  66
             :||: :| :|   |||:|:||:||:::|:  |::|:|  :  |  :  |:
Qy  178  SHKVTLSSWYHDRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQL 237

Db   67  kvym-r-nfkypgdlvlmee-kklnycttgqiwahssylgavfnltvadhlyvnisqlsl 123
             :||: :| ||    ||    |  :|| : |:|| :|:| ||:| : :|| :|
Qy  238  MVYVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSL 297

Db  124  infeesktffglykl 138
             :|| :|:
Qy  298  LDPDQDATYFGAFKV 312
```

```
RESULT 10
ID  R88308 standard; Protein; 178 AA.
AC  R88308;
DT  28-FEB-1996 (first entry)
DE  C-terminally deleted human Fas ligand CD179.
KW  Fas ligand; Tumour Necrosis factor family; apoptosis; cell death;
KW  Fas cell surface antigen; human; Fas-L; CD179; truncated.
OS  Homo sapiens.
PN  WO9513293-A1.
PD  18-MAY-1995.
PF  10-NOV-1994; JP-305975.
PR  10-NOV-1993; JP-305975.
PR  13-DEC-1993; JP-342526.
PR  18-MAR-1994; JP-074344.
PR  08-JUL-1994; JP-180955.
PR  07-SEP-1994; JP-239363.
PR  18-OCT-1994; JP-278378.
PA  (MOCH ) MOCHIDA PHARM CO LTD.
PA  (OSAB-) OSAKA BIOSCIENCE INST.
PI  Nagata S, Nakamura N, Suda T, Takahashi T;
DR  WPI; 95-194031/25.
PT  Peptide which binds to Fas antigen, and antibody reactive with it  -
PT  for treatment and diagnosis of viral or auto:immune diseases
PS  Example 21; Page 196-197; 300pp; Japanese.
CC  CD179 is a truncated human Fas ligand from which the Leu residue at
CC  the C-terminus (position 179) has been deleted. Fas ligands able to
CC  induce apoptosis in cells which express the Fas cell surface antigen
CC  are claimed.
SQ  sequence 178 AA;

Query Match       7.68;  Score 174;  DB 14;  Length 178;
Best Local Similarity 24.08;  Pred. No. 1.43e-05;
Matches  44;  Conservative  54;  Mismatches  75;  Indels  10;  Gaps  8;
```

```
Db    5  lqkelaelr-estsqmhtassle-kqighpspppekkelrkvahtgksnsrsmplewed  62
             ||  :|   :::| :|:|| |  |  |::| | |:  | : | :    :|
Qy  130  LQHIVGPQRFSGAPAMMEGSWLDVAQRGRKPEAQPFAHLTINAASIPSGSHKVTLS-SWYH 188

Db   63  tygivllsgvkykgglvinetglyfvyskvyfrgq-scnnlp---lshkvym-rns-ky 116
             | :: |:  |  :|  |:|| |:::|:||::|:| ||:|   : |:  |:|
Qy  189  DRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYVVKTSIKI 248

Db  117  pqdl-vmmegkmmsycttgqmwarssylgavfnltsadhlyvnveslslvnfeesqtffg 175
             |::| : :|   :|| : : | |:|:| ||:|  | ::|  : :||   :|
Qy  249  PSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDPDQDATYFG 308

Db  176  lyk 178
             :|
Qy  309  AFK 311
```

```
RESULT 11
ID  W11814 standard; peptide; 179 AA.
AC  W11814;
DT  20-OCT-1997 (first entry)
DE  Fas ligand.
KW  Human; Fas ligand; antigen; neutral; antibody; apoptosis; HIV;
KW  induction; assay; enzyme linked immunosorbant assay; diagnosis;
KW  disease; hepatitis B; hepatitis C; human immunodeficiency virus;
KW  graft versus host disease; ulcerative colitis; sequelae;
KW  myocardial infarction; treatment.
OS  Homo sapiens.
FH  Key             Location/Qualifiers
FT  peptide         24..43
FT                  /label= M52
FT  peptide         47..60
FT                  /label= M53
FT  peptide         60..72
FT                  /label= M54
FT  peptide         79..89
FT                  /label= M55
FT  peptide         89..107
FT                  /label= M56
FT  peptide         92..125
FT                  /label= M57
FT  peptide         126..141
FT                  /label= M58
FT  peptide         164..174
FT                  /label= M59
PN  WO9702290-A1.
PD  23-JAN-1997.
PF  01-JUL-1996; J01820.
PR  17-MAY-1996; US-649100.
PR  30-JUN-1995; JP-188480.
PA  (MOCH ) MOCHIDA PHARM CO LTD.
PA  (OSAB-) OSAKA BIOSCIENCE INST.
PI  Co MS, Matsusue T, Nagata S, Shirakawa K, Vasquez M;
DR  WPI; 97-108917/10.
PT  Antibody reactive with Fas ligand capable of inducing apoptosis -
PT  used for diagnostic assay of Fas ligand in body fluids and for
PT  treatment of diseases in which Fas ligand/Fas antigen is involved
PS  Disclosure: Fig 1; 164pp; Japanese.
CC  Antigens derived from the present human Fas ligand sequence can
CC  be used in the preparation of a neutral antibody against an
CC  apoptosis inducing Fas ligand. The antibody can be used in a Fas
CC  ligand assay, e.g. an enzyme linked immunosorbant assay, to
CC  diagnose diseases in which the Fas ligand/Fas antigen system is
CC  implicated, e.g. hepatitis B/C, human immunodeficiency virus,
CC  graft/host disorders, ulcerative colitis or sequelae of myocardial
CC  infarction. The antibody may also be used to treat such diseases.
SQ  sequence 179 AA;

Query Match       7.68;  Score 175;  DB 23;  Length 179;
Best Local Similarity 23.98;  Pred. No. 1.18e-05;
Matches  44;  Conservative  55;  Mismatches  75;  Indels  10;  Gaps  8;
```

```
Db     5 lqkelaelr-estsqmhtassle-kqighpsppekkelrkvahltgksnsrsmplewed 62
         ||  :  |: :  ||: :      | ::  ||  ::   | ::  ::   : ::  :
Qy   130 LQHIVGPQRFSGAPAMMEGSWLDVAQRGKPEAQPFAHLTINAASIPSGSHKVTLS-SWYH 188

Db    63 tygivllsgvkykkgglvinetglyfvyskvyfrgq-scnnlp---lshkvym-rns-ky 116
         |:  ||  |   :  ::  |  ||||| : |  |   : :|    |
Qy   189 DRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYVVKTSIKI 248

Db   117 pqdl-vmmegkmmsycttgqmwarssylgavfnltsadhlyvnvselslvnfeesqttfg 175
         : ||   :  |: | ||   :  ||   || :|:: ||| : || |||||||||||||
Qy   249 PSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDPDQDATYFG 308

Db   176 lykl 179
         |:|
Qy   309 AFKV 312
```

RESULT 12
ID   R79099 standard; Protein; 179 AA.
AC   R79099;
DT   21-FEB-1996 (first entry)
DE   Human Fas ligand (partial sequence).
KW   Fas ligand; Tumour Necrosis factor family; apoptosis; cell death;
KW   Fas cell surface antigen; Fas-L; human.
OS   Homo sapiens.
PN   WO9513293-A1.
PD   18-MAY-1995.
PF   10-NOV-1994; J01899.
PR   10-NOV-1993; JP-305975.
PR   13-DEC-1993; JP-342526.
PR   18-MAR-1994; JP-074344.
PR   08-JUL-1994; JP-180955.
PR   07-SEP-1994; JP-239363.
PR   18-OCT-1994; JP-278378.
PA   (MOCH) MOCHIDA PHARM CO LTD.
PA   (OSAB-) OSAKA BIOSCIENCE INST.
PI   Nagata S, Nakamura N, Suda T, Takahashi T;
     WPI; 95-194031/25.
DR   N-PSDB; Q94157.
PT   Peptide which binds to Fas antigen, and antibody reactive with it -
PT   for treatment and diagnosis of viral or auto:immune diseases
PS   Claim 3; Page 222-224; 300pp; Japanese.
CC   Fas ligands or active fragments able to induce apoptosis in cells
CC   which express the Fas cell surface antigen are claimed. The
CC   proteins are isolated from human, rat and mouse sources. The present
CC   sequence represents part of the human Fas ligand.
SQ   Sequence 179 AA;

Query Match   7.6%;   Score 175;   DB 14;   Length 179;
Best Local Similarity 23.9%;   Pred. No. 1.18e-05;
Matches 44; Conservative 55; Mismatches 75; Indels 10; Gaps 8;

```
Db     5 lqkelaelr-estsqmhtassle-kqighpsppekkelrkvahltgksnsrsmplewed 62
         ||  :  |: :  ||: :      | ::  ||  ::   | ::  ::   : ::  :
Qy   130 LQHIVGPQRFSGAPAMMEGSWLDVAQRGKPEAQPFAHLTINAASIPSGSHKVTLS-SWYH 188

Db    63 tygivllsgvkykkgglvinetglyfvyskvyfrgq-scnnlp---lshkvym-rns-ky 116
         |:  ||  |   :  ::  |  ||||| : |  |   : :|    |
Qy   189 DRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMVYVVKTSIKI 248

Db   117 pqdl-vmmegkmmsycttgqmwarssylgavfnltsadhlyvnvselslvnfeesqttfg 175
         : ||   :  |: | ||   :  ||   || :|:: ||| : || |||||||||||||
Qy   249 PSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDPDQDATYFG 308

Db   176 lykl 179
         |:|
Qy   309 AFKV 312
```

RESULT 13
ID   R79066 standard; Protein; 179 AA.
AC   R79066;
DT   22-FEB-1996 (first entry)

DE   Rat Fas ligand (partial sequence).
KW   Fas ligand; Tumour Necrosis factor family; apoptosis; cell death;
KW   Fas cell surface antigen; Fas-L; rat.
OS   Rattus rattus.
PN   WO9513293-A1.
PD   18-MAY-1995.
PF   10-NOV-1994; J01899.
PR   10-NOV-1993; JP-305975.
PR   13-DEC-1993; JP-342526.
PR   18-MAR-1994; JP-074344.
PR   08-JUL-1994; JP-180955.
PR   07-SEP-1994; JP-239363.
PR   18-OCT-1994; JP-278378.
PA   (MOCH) MOCHIDA PHARM CO LTD.
PA   (OSAB-) OSAKA BIOSCIENCE INST.
PI   Nagata S, Nakamura N, Suda T, Takahashi T;
     WPI; 95-194031/25.
DR   N-PSDB; Q99495.
PT   Peptide which binds to Fas antigen, and antibody reactive with it -
PT   for treatment and diagnosis of viral or auto:immune diseases
PS   Claim 7; Page 215-216; 300pp; Japanese.
CC   Fas ligands or active fragments able to induce apoptosis in cells
CC   which express the Fas cell surface antigen are claimed. The
CC   proteins are isolated from human, rat and mouse sources. The present
CC   sequence represents part of the rat Fas ligand.
SQ   Sequence 179 AA;

Query Match   7.6%;   Score 174;   DB 14;   Length 179;
Best Local Similarity 26.7%;   Pred. No. 1.43e-05;
Matches 36; Conservative 38; Mismatches 53; Indels 8; Gaps 6;

```
Db    53 srsiple-wedtygtalisgvkykkgglvineaglyfvyskvyfrgq-scnsqp---lsh 107
         :    |   |   |:  ||  |   :  ::  |  ||||| : |  |   : :|
Qy   178 SHKVTLSSWYHDRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQL 237

Db   108 kvym-r-nfkypgdlvlmee-kklnycttgqiwahssylgavfnltvadhlyvnisqlsl 164
         :||   :      |: | ||    |   :  ||   || :|:: ||| : ||
Qy   238 MVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSL 297

Db   165 lnfeesktffglykl 179
         |||||||    |:|
Qy   298 LDPDQDATYFGAFKV 312
```

RESULT 14
ID   R79095 standard; Protein; 278 AA.
AC   R79095;
DT   20-FEB-1996 (first entry)
DE   Rat Fas ligand encoded by cDNA clone pTN24-15.
KW   Fas ligand; Tumour Necrosis factor family; apoptosis; cell death;
KW   Fas cell surface antigen; rat; Fas-L.
OS   Rattus rattus.
FH   Key   Location/Qualifiers
FT   region   77..99
FT   /label= transmembrane_anchor
FT   modified_site   116
FT   /label= N-glycosylation_site
FT   modified_site   130
FT   /label= N-glycosylation_site
FT   modified_site   247
FT   /label= N-glycosylation_site
FT   modified_site   257
FT   /label= N-glycosylation_site
PN   WO9513293-A1.
PD   18-MAY-1995.
PF   10-NOV-1994; J01899.
PR   10-NOV-1993; JP-305975.
PR   13-DEC-1993; JP-342526.
PR   18-MAR-1994; JP-074344.
PR   08-JUL-1994; JP-180955.
PR   07-SEP-1994; JP-239363.
PR   18-OCT-1994; JP-278378.
PA   (MOCH) MOCHIDA PHARM CO LTD.

```
PA   (OSAB-) OSAKA BIOSCIENCE INST.
PI   Nagata S, Nakamura N, Suda T, Takahashi T;
DR   WPI: 95-194031/25.
DR   N-PSDB: Q94153, TO3656.
PT   Peptide which binds to Fas antigen, and antibody reactive with it -
PT   for treatment and diagnosis of viral or auto:immune diseases
PS   Claim 8; Fig 2 and 3; 300pp; Japanese.
CC   Clones contg. rat cDNA fragments that encoded Fas ligand-like
CC   proteins were identified by antibody panning. Clone pTN24-15 was
CC   isolated and sequenced (Q94153). The sequence includes an open
CC   reading frame coding for a predicted amino acid sequence of
CC   278 residues. The deduced protein sequence (R79095) contains a
CC   transmembrane anchor region and 4 consensus N-glycosylation sites;
CC   the protein is a claimed Fas ligand able to induce apoptosis in
CC   cells which express the Fas cell surface antigen.
SQ   Sequence    278 AA;

Query Match             7.6%;  Score 174;  DB 14;  Length 278;
Best Local Similarity  26.7%;
Matches   36;  Conservative  38;  Mismatches  53;  Indels   8;  Gaps   6;

       Db   107 lqkelaelr-estsqmhtassle-kqighpsppppekkelrkvahltgksnsrsmplewed 164
                   || :::  :::|  ::||: :|: :|::|::    :|: |::|   |—
       Qy   130 LQHIVGPQRFSGAPAMMEGSWLDVAQRGKPEAQPFAHLTNAASIPSGSHKVTLS-SWYH 188

       Db   165 tygivilsgvkykkgglvineaglyfvyskvyfrgq-scnsqp--lsh 206
                   :|: :|| |::|:|::|:|:::|:  || : :||—| |::|
       Qy   152 srsiple-wedtygtalisgvkykkgglvineaglyfvyskvyfrgq-scnslp--lshkvym-rns-ky 218

       Db   189 DRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQLMYYVVKTSIKI 248
                   |—| :::| ||::||:|:||::  |||| || ||
       Qy   178 SHKVTLSSWYHDRGWAKISNMTLSNGKLRVNQDGFYYLYANICFRHHETSGSVPTDYLQL 237

       Db   207 kvym-r-nfkypqdlvlmee-kklnycttgqiwahssylgavfnltvadhlyvnisqls1 263
                   :|| :::| |:: ||—| | |:::| ::|| ::|: ||:|||||
       Qy   219 pqdl-vmmegkmmnsycttgqmwarssylgavfnltsadhlyvnvselslvnfeesqtffg 277

       Db   238 MVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSL 297
                   :|:  ||   |:|| ::|||||:|||||::||  | |:|::||:|:||||::|::|::|||
       Qy   249 PSSHNLMKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSNPSLLDPDQDATYFG 308

       Db   264 infeesktffglykl 278
                   ::|:|| |:||::||
       Qy   298 LDPDQDATYFGAFKV 312

       Db   278 lykl 281
                   :|:
       Qy   309 AFKV 312

Search completed: Fri Nov 13 17:13:30 1998
Job time : 169 secs.


RESULT   15
ID   W27143 standard; Protein; 281 AA.
AC   W27143;
DT   23-MAR-1998  (first entry)
DE   Human Fas ligand.
KW   Fas ligand fusion protein; FasL; antigen; autoimmune disorder;
KW   cell surface marker; extracellular domain; cancer.
OS   Homo sapiens.
PN   WO9733617-A1.
PD   18-SEP-1997.
PF   11-MAR-1997; U03571.
PR   13-MAR-1996; US-614584.
PA   (PROT-) PROTEIN DESIGN LABS INC.
PA   Queen CL, Schneider WP, Vasquez M;
WPI: 97-470649/43.
N-PSDB: T85236.
PT   Fusion protein containing recognition region of Fas ligand and
PT   specific binding agent - reactive with cell surface marker, and used
PT   for treating autoimmune disease and cancer
PS   Example 7; Page 26; 39pp; English.
CC   This sequence represents the human Fas ligand. The nucleotide
CC   sequence encoding the Fas ligand was used in the construction of a Fas
CC   ligand fusion protein for treatment of autoimmune disease. This new
CC   fusion protein comprises a recognition sequence of the extracellular
CC   domain of a Fas ligand (FasL) protein and a polypeptide that binds
CC   specifically to a cell surface marker. These fusion proteins are used
CC   to treat autoimmune disease (particularly multiple sclerosis, rheumatoid
CC   arthritis, insulin-dependent diabetes mellitus and inflammatory
CC   bowel disease, or cancers (e.g. of breast, lung, colon, ovary or
CC   prostate). They act by suppressing T-cell mediated immune responses and
CC   can also be used to detect Fas-expressing cells in patients with
CC   leukaemia and for destruction of such cells in vitro. Fusion proteins
CC   are easily and economically produced, are required in only small amounts
CC   and have little or no immunogenicity.
SQ   Sequence    281 AA;

Query Match             7.6%;  Score 175;  DB 26;  Length 281;
Best Local Similarity  23.9%;
Matches   44;  Conservative  55;  Mismatches  75;  Indels  10;  Gaps   8;
```

```
*********************************************************
*                                                       *
*   M P S R C H                                 (TM)    *
*                                                       *
*********************************************************
```

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
           Distribution rights by Oxford Molecular Ltd

MPsrch_nn   n.a. - n.a. database search, using Smith-Waterman algorithm

...on on: Sat Nov 14 02:08:59 1998;   MasPar time 2437.70 Seconds
                                         1379.049 Million cell updates/sec
...abular output not generated.

Title: >US-08-989-362-1
Description: (1-2191) from US08989362.seq
Perfect Score: 2191
N.A. Sequence: 1 GCCAGGACCTCTGTGAACCG.........TTTTGTGTACTTAAAAATGGC 2191
     Comp: CGGTCCTGGAGACACTTGGC.........AAAACCATGAATTTTTACCG

Scoring table: TABLE default
                Gap 6

Nmatch STD :   Dbase 0; Query 0

Searched: 1988617 seqs, 767163441 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database: embl-est55
             1:em_est10 2:em_est11
Database: genbank-est107
             3:gb_est1 4:gb_est10 5:gb_est11 6:gb_est12 7:gb_est13
             8:gb_est14 9:gb_est15 10:gb_est16 11:gb_est17
             12:gb_est18 13:gb_est19 14:gb_est2 15:gb_est20
             16:gb_est21 17:gb_est22 18:gb_est23 19:gb_est24
             20:gb_est3 21:gb_est4 22:gb_est5 23:gb_est6 24:gb_est7
             25:gb_est8 26:gb_est9 27:gb_gss

...atistics: Mean 12.352; Variance 2.924; scale 4.224

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result No.	Score	Query Match %	Length	DB	ID	Description	Pred. No.
	1	518	23.6	530	26	AA170348	ms88g11.r1 Soares mous	0.00e+00
	2	133	6.1	419	8	AA504450	aa59h12.r1 NCI_CGAP_GC	7.17e-164
	3	132	6.0	375	8	AA504211	aa59h12.s1 NCI_CGAP_GC	3.33e-162
c	4	64	2.9	252	13	AA754211	97SN1787 Rice Immature	1.73e-53
c	5	54	2.5	247	13	AA754458	97SN1784 Rice Immature	6.74e-39
c	6	55	2.5	252	11	AA754459	97SN1787 Rice Immature	2.51e-40
	7	55	2.5	2275	11	AF034173	Homo sapiens ntcon2 co	2.51e-40
c	8	47	2.1	2275	11	AF034173	Homo sapiens ntcon2 co	4.02e-29
c	9	40	1.8	247	9	AA754458	97SN1784 Rice Immature	7.75e-20
c	10	30	1.4	282	9	AA573470	nm53f09.s1 NCI_CGAP_Br	8.08e-08
c	11	31	1.4	707	27	B19588	T2203-T7.1 TAMU Arabid	6.20e-09
c	12	28	1.3	189	27	AQ068616	HS_2249_A1_H12_MF CIT	1.16e-05
c	13	29	1.3	407	27	AQ022111	CIT-HSP-2303D3.TR CIT-	9.96e-07
	14	29	1.3	419	27	AQ046443	RPC111-34J1.TK RPC111	9.96e-07
	15	28	1.3	488	18	AI041106	ov61f03.x1 Soares_test	1.16e-05
c	16	29	1.3	623	27	AQ000570	CIT-HSP-228007.TF CIT-	9.96e-07
	17	28	1.3	660	11	AF034177	Homo sapiens ntcon6 co	1.16e-05
	18	28	1.3	719	27	B21675	F17P15-Sp6 IGF Arabido	1.16e-05
c	19	26	1.2	84	19	AA966799	s9d02a1.r1 Aspergillus	1.29e-03
	20	26	1.2	182	16	C89935	Dictyostelium discoide	1.29e-03
c	21	27	1.2	254	25	AA204255	mu26g12.r1 Soares 2NDM	1.26e-04
	22	26	1.2	273	18	AA257887	EST 372 Boophilus micr	1.29e-03
	23	26	1.2	280	16	AA832369	oc92b09.s1 NCI_CGAP_GC	1.29e-03
	24	26	1.2	294	25	AA077530	7B44H12 Chromosome 7 F	1.29e-03
	25	26	1.2	300	7	AA550088	ng40a02.s1 NCI_CGAP_Co	1.29e-03
c	26	26	1.2	312	8	CELK13G6R	118m3 gmbPfHB3.1, G.	1.29e-03
c	27	27	1.2	337	7	C23862	C.elegans cDNA clone y	1.26e-04
	28	26	1.2	339	24	N98120	Dictyostelium discoide	1.29e-03
c	29	26	1.2	355	22	AA230383	221C3 czapPFDd2.1, De	1.29e-03
	30	26	1.2	363	26	AA938196	mv73a12.r1 Soares mous	1.29e-03
c	31	26	1.2	402	17	AA255819	oc10e12.s1 NCI_CGAP_GC	1.29e-03
	32	27	1.2	414	26	AA701979	zr84g08.r1 Soares NHHM	1.26e-04
c	33	26	1.2	439	17	AA004008	zi68d12.s1 Soares feta	1.29e-03
c	34	27	1.2	460	23	AA776640	mg80c11.r1 Soares mous	1.26e-04
c	35	26	1.2	460	13	AA425966	ae79g07.s1 stratagene	1.29e-03
	36	27	1.2	463	19	AA948367	zw17a06.r1 Soares ovar	1.26e-04
	37	26	1.2	513	19	AA279926	oq46e12.s1 NCI_CGAP_Ki	1.29e-03
	38	26	1.2	514	8	AA291434	zs92g01.r1 NCI_CGAP_GC	1.29e-03
	39	26	1.2	515	8	N98261	zt44h08.r1 Soares ovar	1.29e-03
	40	26	1.2	550	22	AA428848	za28f09.r1 Homo sapien	1.29e-03
	41	27	1.2	579	10	N58217	zv50h12.r1 Soares ovar	1.29e-03
	42	26	1.2	581	25	AA433286	yv66h06.s1 Soares feta	1.29e-03
c	43	26	1.2	586	16	C91060	EST100 Drosophila mela	1.29e-03
	44	26	1.2	696	16	C90750	Dictyostelium discoide	1.29e-03
	45	27	1.2	866	17		Dictyostelium discoide	1.26e-04

ALIGNMENTS

```
RESULT      1
LOCUS       AA170348        530 bp    mRNA      EST       16-FEB-1997
DEFINITION  ms88g11.r1 Soares mouse 3NbMS Mus musculus cDNA clone 618692 5',
            mRNA sequence.
ACCESSION   AA170348
NID         g1748900
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Mus.
REFERENCE   1  (bases 1 to 530)
AUTHORS     Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE       The WashU-HHMI Mouse EST Project
JOURNAL     Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:379516
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 490.
FEATURES            Location/Qualifiers
     source         1..530
                    /organism="Mus musculus"
                    /strain="C57BL/6J"
```

```
                   /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                   polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                   was primed with a Not I - oligo(dT) primer [5'
                   TGTTACCAATCTGAAGTGGGACGGCGCCGGGCGTGTTTTTTTTTTTTTTTTTTTT
                   3']; double-stranded cDNA was ligated to Eco RI adaptors
                   (Pharmacia), digested with Not I and cloned into the Not I
                   and Eco RI sites of the modified pT7T3 vector. RNA
                   provided by Dr. Bertrand Jordan. Library went through
                   three rounds of normalization, and was constructed by
                   Bento Soares and M.Fatima Bonaldo."
                   /db_xref="taxon:10090"
                   /clone="618692"
                   /clone_lib="Soares mouse 3NbMS"
                   /sex="male"
                   /tissue_type="Spleen"
                   /dev_stage="4 weeks"
                   /lab_host="DH10B"

BASE COUNT      177 a       72 c       85 g      196 t
ORIGIN
        <1. .>530

Query Match          23.6%;  Score 518;  DB 26;  Length 530;
Best Local Similarity 99.2%;  Pred. No. 0.00e+00;
        Matches  527;  Conservative  0;  Mismatches  3;  Indels  1;  Gaps  1;

Db      1  AATGTTTTCTGTGCAAAGTTTGTAAATTATATTGTGCTATAGTATTGATTCAAATA   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1588  AATGTTTTCTGTGCAAAGTTTGTAAATTATATTGTGCTATAGTATTGATTCAAATA 1647

Db     61  TTTAAAAATGCTCACTGTGTGACATATTTAAATTTAAAATTACAGATGTATTTAACTG  120
           ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Qy   1648  TTTAAAAATGCTCACTGTGTGACATATTTAAATTTAAAATTACAGATGTATTTAACTG 1707

Db    121  GTGCACTTTGTAATTCCCCTGAAGTACTCGTAGCTCAGGGGCAGAGTACTGTTTCTGG  180
           ||||||||||||||||||||||||||||||||||| |||||||||||
Qy   1708  GTGCACTTTGTAATTCCCCTGAAGTACTCGTAGCTCAGGGGCAGAATACTGTTTCTGG 1767

Db    181  TGACCACATGTAGTTTATTTCTTTTATTCTTTTTAACTTAATAGAGTCTTCAGACTTGTCA  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1768  TGACCACATGTAGTTTATTTCTTTTATTCTTTTTAACTTAATAGAGTCTTCAGACTTGTCA 1827

Db    241  AAACTATGCAAGCAAGATAATAAATAAATAAAATAAAATGAATACCTTGAATAATAAGTAGG  300
           |||||||||||||||||||||||||||||||||||||| |||||||||||||
Qy   1828  AAACTATGCAAGCAAGATAATAAATAAAATAAAATGAATATCTTGAATAATAAGTAGG 1887

Db    301  ATGTTGGTCACCAGGTGCCTTCAAATTTAGGAGCTAATTGACTTTAGGAGCTGACATAG  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1888  ATGTTGGTCACCAGGTGCCTTCAAATTTAGGAGCTAATTGACTTTAGGAGCTGACATAG 1947

Db    361  CCAAAAGGATACATATAGGCTACTGAAA-TCTGTCAGGAGTATTTATGCAATTATTGA  419
           ||||||||||||||||||||||||||| ||||||||||||||||||||
Qy   1948  CCAAAAAGGATACATATAGGCTACTGAAAATCTGTCAGGAGTATTTATGCAATTATTGA 2007

Db    420  ACAGGTGTCTTTTTTTTACAAGAGCTACAAATTGTAAATTTGTTTGTTTTCTTTTTTCCCATA  479
           ||||||||||||||||||||||||||||||||||||||||||||| | ||||||||
Qy   2008  ACAGGTGTCTTTTTTTTACAAGAGCTACAAATTGTAAATTTGTTTGTTTTCTTTTTTCCCATA 2067

Db    480  GAAAATGTACTATAGTTTATCAGCCAAAAAAACAATCCACTTTTAATTAG  530
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2068  GAAAATGTACTATAGTTTATCAGCCAAAAAAACAATCCACTTTTAAATTAG 2118

RESULT    2
LOCUS       AA504450       419 bp    mRNA          EST       18-AUG-1997
DEFINITION  aa59h12.r1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:825287 5',
            mRNA sequence.
ACCESSION   AA504450
NID         g2240610
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
```

```
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 400.
FEATURES             Location/Qualifiers
     source          1..419
                     /organism="Homo sapiens"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5' TGTTACCAATCTGAAGTGGGACGGCGCCGGGCGTTTTTTTTTTTTTTTTTTTT
                     3']. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
                     /db_xref="taxon:9606"
                     /clone="IMAGE:825287"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
BASE COUNT      121 a       58 c       96 g      144 t
ORIGIN
        <1. .>419

Query Match           6.1%;  Score 133;  DB 8;  Length 419;
Best Local Similarity 83.3%;  Pred. No. 7.17e-164;
        Matches  204;  Conservative  0;  Mismatches  35;  Indels  6;  Gaps  5;

Db    181  CCCTGGTCATGTGCCCCTTCGCAGCT-GAAGTGGAGAGGGTGTCATCT-AGCGCAATTGA  238
           ||||||| |||| || ||| |||| | ||||||| ||  ||| |||| - ||||
Qy   1395  CCCTGGACATGTGCCACTGGACCTTGAAACCTTGAAATTAAGAAGATGCCATTGTCATTGCAAAGAA 1454

Db    239  AGGATCATCTGAAGGGGCAAATTCTTTGAATTGTTACATCATGCTGGAACCTGCAAAAA  298
           |||| |||||||| | |||||||||||||| || ||||||||
Qy   1455  ATGATAGTGTGAAGGGTGAAGGTTAGTTCTTTGAATTGTTACATTGCCGCTGGGACCTGCAAATA 1514

Db    299  A-TACTTTTT--CTAATGAGGAGAGAAAA-TATATGTATTTTTATAATAATCTAAAGTT  354
           | |-||||||  |||||||||||||||||| |||||||||||||||||||||||||||
Qy   1515  AGTTCTTTTTTCTAATGAGGAGAGAGAAAAATATATGTATTTTTTATAATAATGTCTAAAGTT 1574

Db    355  ATATTTCAGATGTAATGTTTTCTTTGCAAAGTATTGTAAAATATATTTGTGCTATAGTAT  414
           |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Qy   1575  ATATTTCAGGTGTAATGTTTTCTGTGCAAAGTTTGTAAAATATATTTGTGCTATAGTAT 1634

Db    415  TTGAT  419
           |||||
Qy   1635  TTGAT 1639

RESULT    3
```

```
LOCUS       AA504211      375 bp    mRNA                 EST       18-AUG-1997
DEFINITION  aa59h12.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:825287 3',
            mRNA sequence.
ACCESSION   AA504211
NID         g2240371
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Mammalia; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 375)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_strausberg@nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image.html

            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 287.
FEATURES    Location/Qualifiers
     source          1..375
                     /organism="Homo sapiens"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCGAAGTGGAGACGGCCGCCTCATTTTTTTTTTTTTTTTTT-
                     3']. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
                     /db_xref="taxon:9606"
                     /clone="IMAGE:825287"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
     mRNA            complement(<1..>375)
BASE COUNT      119 a     59 c     48 g    149 t
ORIGIN

Query Match             6.08%;  Score 132;  DB 8;  Length 375;
  Best Local Similarity   81.4%;  Pred. No. 3.33e-162;
  Matches  275;  Conservative   0;  Mismatches  47;  Indels 16;  Gaps 12;

Db        43 TTTTTGTACCAAAAATTAACATTCAAAGGCAATGTTTTTATTGTACAGTATAAGATAAC 102
             |||| |||||||||| |||||||| |||||||  ||||| |||||||||| ||||||
Cp      2185 TTTAGTACCAAAAAATGC-CATTCAGAAAACAATGCTTTTATTGTACAGTATAATAAAAT 2127

Db       103 A--TTT-ACTAAATTAAAAT-TGGACT--TTTTTG-CTAATAAACCACTAATAACTTTTC 155
             |  |||  |||||| ||||| | ||||  ||||||  | ||| ||| || || |||||
Cp      2126 AACTTTCACTAAATTAAAAAGTGAATTGTTTTTGGCTGATAAACTA-TAGTACATTTTC 2068

Db       156 TATTTGAAAAAC--AA-AAAAGGTACAATTGCGGCACTGTGGAAAA--CACCTGT 210
             ||| |||||||  || |||||| |||| |||||||||| ||||||  |||||||
Cp      2067 TATGGGAAAAAAAAAAGAAAACAAAATTTACAATTGTAGCTCTTGTAAAAAAAGACACCTTT 2008

Db       211 TCAACAATATATAACTACTCTTGACAGATTT-CAGTGGCCCATTATGTATCCTTTTGG 269
             |||||| ||| |||| ||||| ||| |||| ||||||||||||||||||||||||||
Cp      2007 TCAATAATTGCATAAATACTCCTGACAGATTTCAGTAGCCTATTATGTATCCTTTTGG 1948

Db       270 CAATGTCAGCTTCTCAAAGTCAATTAGTTCTCTAAATTGAAAGGCACCTGGTGGCCAACA 329
             ||||||||| |||||||||||| |||| |||||||||||||||||||||||| |||||
Cp      1947 CTATGTCAGCTC-CTAAAGTCAATTAGCGTCTCTAAATTTGAAAGGCACCTGGTGACCAACA 1889

Db       330 TCCTGCTTATTATTCAAGGCATCCATTTAATTTTTT 367
             ||| ||||||||||||||||  ||||||| |||| ||
Cp      1888 TCCTACTTATTATTCAAGATATTCATTTTATTTTATT 1851


RESULT 4
LOCUS       AA754459       252 bp    mRNA                           20-JAN-1998
DEFINITION  97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
            cDNA clone 97SN1787, mRNA sequence.
ACCESSION   AA754459
NID         g2801165
KEYWORDS    EST.
SOURCE      rice.
ORGANISM    Oryza sativa
            Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
            Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
            Poales; Poaceae; Oryza.
REFERENCE   1  (bases 1 to 252)
AUTHORS     Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R, Moon,E.P.,
            Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
            Lee,M.C. and Eun,M.Y.
TITLE       Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL     Unpublished (1998)
COMMENT     Contact: Eun M.Y.
            Department of Cytogenetics
            National Inst. of Agri. Sci. and Tech., RDA
            Suwon, Kyunggido, Korea
            Tel: 82 331 290 0301
            Fax: 82 331 290 0307
            Email: myeun@sun20.asti.re.kr
            Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
            University, Yongin, Korea. 449-728 bhnahm@bioserver.myongji.ac.kr
            Seq primer: M13 Reverse Primer
FEATURES    Location/Qualifiers
     source          1..252
                     /organism="Oryza sativa"
                     /cultivar="Milyang23"
                     /note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
                     XhoI; Directional cDNA library inserted into lambda zAPII
                     vector at 5'end with EcoRI and 3' end with Xho I site."
                     /db_xref="taxon:4530"
                     /clone="97SN1787"
                     /clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
                     /tissue_type="Immature Seed"
                     /dev_stage="5 days after pollination"
                     /lab_host="E. coli SOLR"
BASE COUNT       5 a     21 c     12 g     35 t    179 others
ORIGIN

Query Match             2.9%;  Score 64;  DB 13;  Length 252;
  Best Local Similarity   10.7%;  Pred. No. 1.73e-53;
  Matches   24;  Conservative 117;  Mismatches  82;  Indels 2;  Gaps  2;

Db        23 YBCHGNBVWVCVASHGNYMSVHNCTBRGTHCDCKNVNWSTMTWGTVNWBNVSGDWHYWBV 82
             :|: || :||:::||::|: |||  |: || ||::::|: :|:|::|::: ||::| :|:
Cp       839 CCATCAGCTGAAGATAGTCTGTAGGTACCGTTCCCGATGTTTCATGATGCCGAAGCAAA 780

Db        83 BNTKVDVGNHTRCSRWRBVTRMAHYHDYTNCBBYNNNDYHMWHBBMYBBTGCMTCMWCW 142
             : |::::||:|:|::| |:|:|||: ::: ::| ::::: :: :|:  |: ::|:|:| |
Cp       779 TGTTGGCGTACAGGTAATGAGAAGCCATCTTGGTTAACCCT-TAGTTTTCCGTTGCTTAAC 721

Db       143 BHYNTKCTASGWHTSTNFDVKSSTNTWGVTBSYDKSMHGYWCSBBVKYHTKVSTTRATRS 202
             ::: :| ::|::|: |::|:|||| : :|: ::::|: :|| :|:::: |:|:::::|: |
Cp       720 GTCATGTTAGAGATCTTGGCCCAGCCTCGATCGTGGTACCAAGAGGACA-GAGTGACTTT 662
```

Left column:

```
Db   203  YTCVRKYCVMWMTKKVVKKYHVBBGCHBTDSKCKTMWMTNKHVM 247
          |:|| :|::::::::::::::||:|| :| :| ::|:
Cp   661  ATGGGAACCCGATGGGATGCTGGCAGCATTGGTGAGGTGTGC 617

RESULT    5
LOCUS         AA754458       247 bp       mRNA         EST      20-JAN-1998
DEFINITION    97SN1784 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
              cDNA clone 97SN1784, mRNA sequence.
ACCESSION     AA754458
NID           g2801164
KEYWORDS      EST.
SOURCE        rice.
ORGANISM      Oryza sativa
              Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
              Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
              Poales; Poaceae; Oryza.
REFERENCE     1  (bases 1 to 247)
AUTHORS       Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
              Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
              Lee,M.C. and Eun,M.Y.
TITLE         Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL       Unpublished (1998)
COMMENT       Contact: Eun M.Y.
              Department of Cytogenetics
              National Inst. of Agri. Sci. and Tech, RDA
              Suwon, Kyunggido, Korea
              Tel: 82 331 290 0301
              Fax: 82 331 290 0307
              Email: myeun@sun20.asti.re.kr
              Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
              University, Yongin, Korea. 449-728 bhnahm@bioserver.myongji.ac.kr
              Seq primer: M13 Reverse Primer.
FEATURES      Location/Qualifiers
source        1..247
              /organism="Oryza sativa"
              /cultivar="Milyang23"
              /note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
              XhoI; Directional cDNA library inserted into lambda ZAPII
              vector at 5'end with EcoRI and 3' end with Xho I site."
              /db_xref="taxon:4530"
              /clone="97SN1784"
              /clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
              /tissue_type="Immature Seed"
              /dev_stage="5 days after pollination"
              /lab_host="E. coli SOLR"
BASE COUNT       7 a      16 c      21 g      34 t      169 others
ORIGIN

Query Match      2.5%;      Score 54;      DB 13;      Length 247;
Best Local Similarity 11.3%;  Pred. No. 6.74e-39;
Matches   26;  Conservative  109;  Mismatches  94;  Indels  1;  Gaps  1;

Db    17  BAWMNKHTHMMTBBWCCVRRVGTTNNGKHNGRTTWNDCSDNAHCRYTVBWYYARSKYG 76
          :|: |:| ::|| ::|| |:|| :| |: :| :||: |::|  ||:  |||:
Cp  1496  CAATGTAACAATTCAAAAGAACTTAACCCTTCACACTATCATTTCTTTGCAATGACATGG 1437

Db    77  YGTBYYSWNVDTNTGGTGVGKTTVNVHSGWNNRCSNSVVYVWBTAYCDYBHYBDRANHVD 136
          :|::::::: :|| |  || || ::||| ||:: :  |:|| ::: ||
Cp  1436  CATCTTCTTAATTCAAGGTTCTCAGTGGCACATGTCCAGGGGTTAGACCCGAGACTGT 1377

Db   137  DTRCTNDRGYCNYTASDNGTSATKRVTGYDKIDSDCGGGCW-RKVTYGSSBYBRCGVNVM 195
          : |:::| :| :: :::  :| :|  ||: |:|  :|
Cp  1376  GACCCCCTTCCATAGCTCACGTGTAAGTTTTCATAAGCATCGGAATACCTCTCCCAATC 1317

Db   196  VRTTSMWTDKSTKMBSMDMSRSRVHYGRWMBNKKRGMSRNWTDTKTWRT 245
          :|| |::::||  : |:  ::| : |  |: ::| :::|::||:|:|
Cp  1316  TGGTTCAATTCTTAGGAAATCATTACAAAAATTGTAAAACCGGTTGTGTAAT 1267

RESULT    6
```

Right column:

```
LOCUS         AA754459       252 bp       mRNA         EST      20-JAN-1998
DEFINITION    97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
              cDNA clone 97SN1787, mRNA sequence.
ACCESSION     AA754459
NID           g2801165
KEYWORDS      EST.
SOURCE        rice.
ORGANISM      Oryza sativa
              Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
              Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
              Poales; Poaceae; Oryza.
REFERENCE     1  (bases 1 to 252)
AUTHORS       Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
              Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
              Lee,M.C. and Eun,M.Y.
TITLE         Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL       Unpublished (1998)
COMMENT       Contact: Eun M.Y.
              Department of Cytogenetics
              National Inst. of Agri. Sci. and Tech, RDA
              Suwon, Kyunggido, Korea
              Tel: 82 331 290 0301
              Fax: 82 331 290 0307
              Email: myeun@sun20.asti.re.kr
              Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
              University, Yongin, Korea. 449-728 bhnahm@bioserver.myongji.ac.kr
              Seq primer: M13 Reverse Primer.
FEATURES      Location/Qualifiers
source        1..252
              /organism="Oryza sativa"
              /cultivar="Milyang23"
              /note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
              XhoI; Directional cDNA library inserted into lambda ZAPII
              vector at 5'end with EcoRI and 3' end with Xho I site."
              /db_xref="taxon:4530"
              /clone="97SN1787"
              /clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
              /tissue_type="Immature Seed"
              /dev_stage="5 days after pollination"
              /lab_host="E. coli SOLR"
BASE COUNT       5 a      21 c      12 g      35 t      179 others
ORIGIN

Query Match      2.5%;      Score 55;      DB 13;      Length 252;
Best Local Similarity 12.7%;  Pred. No. 2.51e-40;
Matches   30;  Conservative  117;  Mismatches  85;  Indels  4;  Gaps  4;

Db    18  WMTTSYBCHGNBVWVCVASHGNYMSVHNCTBRGTHCDCKNVNWSTMTWGTVNWBNVSGDW 77
          :||||:||| :: :| : |:| :|::| ::| : ||||:  ::|::
Qy  1484  AATTGTTACATTGCGCTGGGACCTGCAAATAAGTTCTTTTTTCTAATGAGGAGAGAAA 1543

Db    78  HYWBVBNTKVDVGNHTRCSRWRBVTRMAHYHDYTNCBBYNNNDYHWHBBMYBBTGCMTC 137
          :|| : |::: | : ||||| ::: ::| | | ::
Qy  1544  ATATATGTATTTTATATAATGTCTAAAGTTATAT-TTCAGGTGTGAATGTTTTCTGTG-C 1601

Db   138  TMWCWBHYNTKCTASGWHTSSTNYDVKSSTNWGVTBSYDKSMHGYWCSBBVKYHTKVSTT 197
          :: ::|:  ||  | |  ||  |:  |: :::|:
Qy  1602  AAAGTTTGTAAATTATATTTGTGCTATAGTATTTGATTCAAAATATTAAAAATGTCTC 1661

Db   198  RATRSYTCVRKYCVMWMTKKVVKKYHVVBBGCHBTDSKCKT-MWMTNKHVMTSTTD 252
          :|| ::| | ||| : :||:|  :|   |::||| |::|:| :| |::||:
Qy  1662  ACTGTTGACATATTTAATGTT-TRAAATGTACAGATGTATTTAACTGCTGGTGCACTTT 1716

RESULT    7
LOCUS         AF034173       2275 bp      mRNA                  22-DEC-1997
DEFINITION    Homo sapiens ntcon2 contig mRNA, partial sequence, mRNA sequence.
ACCESSION     AF034173
NID           g2707735
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
```

RESULT 8
LOCUS/DEFINITION AF034173 2275 bp mRNA EST 22-DEC-1997
DEFINITION Homo sapiens ntcon2 contig mRNA, partial sequence, mRNA sequence.
ACCESSION AF034173
NID g2707735
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
          Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2275)
AUTHORS Tripodis,N. and Ragoussis,J.
TITLE Generation of a transcription map in the region immediately
       centromeric to human MHC across the 6p21.2-6p21.3 chromosomal
       boundary
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2275)
AUTHORS Tripodis,N. and Ragoussis,J.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-1997) Division of Medical and Molecular Genetics,
         Guy's Hospital, 7th floor, Guy's Tower, London SE1 9RT, UK
FEATURES Location/Qualifiers
source 1..2275
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="6"
        /map="6p21.3"
        /clone="ntcon2 contig"
        /tissue_type="liver; brain"
        /dev_stage="fetus"
        /note="similar to Br140"

BASE COUNT 438 a 619 c 470 g 599 t 149 others
ORIGIN

Query Match   2.5%; Score 55; DB 11; Length 2275;
Best Local Similarity 15.4%; Pred. No. 2.51e-40;
Matches 28; Conservative 98; Mismatches 53; Indels 3; Gaps 3;

Db 1477 AMWRYKRWKRKRKGRRKRMTGMYKRMYRAMMMAMCAMMACWWYYWKMRGMKKCWKYRKYKK 1536
          |:||| ::::: ||||| :::|||| :: ||||| :: :||::: :|::
Qy 1642 AAAATATTTAAAAAATGCTCACTGTTGACATATTTAAGTTTAAATG-TACAGATGTAT 1700

Db 1537 YTSTYYKSWSRWYWTYTYWYCWCCTSMKSASCAMMRWMGYMGSRSSRSYWGYWGSMSG 1596
          :| :::::::|::: ||:: |::|| :::: |||| :::::
Qy 1701 TTAACTGGTGCACTTGTRAATTCCCCTGAAGGTACTCGTAGCTAAGGGGCCAGAATACTG 1760

Db 1597 CYGMTKRYRYRYSWTGWTKWTTWWTIWWSMTRWTMTYYTWWTWRTTKTWWWWWTTCWTMR 1656
          : :|:||: :::: :: :| :: :|:: :||::::::|:::||:||:| :: ||:|:::
Qy 1761 TTTCTGGTGACCACGTAGTTTATTTCT-TTAATCTTTTTAACTT-AATAGAGTCTTCA 1818

Db 1657 KR 1658
          :: ::
Qy 1819 GA 1820

RESULT 9
LOCUS AA754458 247 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1784 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
            cDNA clone 97SN1784, mRNA sequence.
ACCESSION AA754458
NID g2801164
KEYWORDS EST.
SOURCE rice.
ORGANISM Oryza sativa
          Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
          Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
          Poales; Poaceae; Oryza.
REFERENCE 1 (bases 1 to 247)
AUTHORS Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R, Moon,E.P.,
         Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
         Lee,M.C. and Eun,M.Y.
TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL Unpublished (1998)
COMMENT Contact: Eun M.Y.
         Department of Cytogenetics
         National Inst. of Agri. Sci. and Tech., RDA
         Suwon, Kyunggido, Korea
         Tel: 82 331 290 0301
         Fax: 82 331 290 0307
         Email: myeun@sun20.asti.re.kr
         Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
         University, Yongin, Korea. 449-728 bnnahm@bioserver.myongji.ac.kr
         Seq primer: M13 Reverse Primer
FEATURES Location/Qualifiers
source 1..247
        /organism="Oryza sativa"
        /cultivar="Milyang23"
        /note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
        XhoI; Directional cDNA library inserted into lambda ZAPII
        vector at 5'end with EcoRI and 3' end with Xho I site."
        /db_xref="taxon:4530"
        /clone="97SN1784"
        /clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
        /tissue_type="Immature Seed"
        /dev_stage="5 days after pollination"
        /lab_host="E. coli SOLR"

BASE COUNT 7 a 16 c 21 g 34 t 169 others
ORIGIN

/map="6p21.3"
/clone="ntcon2 contig"
/tissue_type="liver; brain"
/dev_stage="fetus"
/note="similar to Br140"

BASE COUNT 438 a 619 c 470 g 599 t 149 others
ORIGIN

Query Match   2.1%; Score 47; DB 11; Length 2275;
Best Local Similarity 14.9%; Pred. No. 4.02e-29;
Matches 28; Conservative 92; Mismatches 67; Indels 1; Gaps 1;

Db 1472 AAAAAAMWRYKRWKRKRKGRRKRMTGMYKRMYRAMMMAMCAMMACWWYYWKMRGMKKCWKY 1531
          |||||| :::: :: |||||: :: :|: :| :::: :: :|||:: :| :|
Cp 2023 AAAAAAGACACCTGTTCAATAATTGCATAAATACTCCTGACAGATTTCAGTAGCCTAT 1964

Db 1532 RKYKKYTSTYYKSWSRWYWTYTYT-YWYCWCCTSMKSASCAMMRWMGYMGSRSSRSYWGY 1590
          :|:: :: ::::: :::: |::::::|: :: :::::: :::: :| :: :|
Cp 1963 TATGTATCCTTTTGGCTATGTCAGCTCCTAAAGTCAATTAGCTTCTAAATTGAAAGGC 1904

Db 1591 WGSMSGCYGMTKRYRYRYSWTGWTKWTTWWTWWSMTRWTMTYTWWTWRTTKTWWWWWT 1650
          :|:: :| :: :: :: :||:||:| :||::::|:::|::||:|:: :|:||:|
Cp 1903 ACCTGGTGACCAACATCCTACTTATTATTCAAGATATTCATTTTATTTTATTTATTTAT 1844

Db 1651 TCWTMRKR 1658
          :| ::::
Cp 1843 TTTGCTTG 1836

Query Match   1.8%; Score 40; DB 13; Length 247;
Best Local Similarity 10.1%; Pred. No. 7.75e-20;
Matches 17; Conservative 89; Mismatches 60; Indels 3; Gaps 3;

```
Db   69 YYARSKYGYGTBYSWNVDTNTGGYGVGKTTVNVHSGWNNRCSNSVVYVWBTAYCDYBHY 128
          ::::: :: :: :: :::: :: ::: : :::::: :::: ::::::::::
Qy    2 CCAGGACCTCTGTGAACCGGTCGGGGGGGGGGCCGCCTGGCCGGAGTCTCGCGGCGGT 61

Db  129 BDRANHVDDTRCTNDRGYCNYTASDNGTSATKRVTGVYDKTDSDCGGGCWRKVTYGSSBYB 188
          :::: :: :::: :: ::: ::::: ::: ::: :: :: :::::::: ::::
Qy   62 GGGTGGCCGAGGAAGGGAGGAGAACGA-TCGGCGGA-GCAGGGGCGCCCCAACTCCGGGCGCC 119

Db  189 RCGVNVMVRRTTSMWTDKSTKMBSMDMSRRSRVHYGRWMBNKKRGMSRNW 237
          ::: :  ::: ::::::: :: :: ::::::::
Qy  120 GCGC-CATGCGCCCGGGCCAGCCGAGACTACGGCAAGTACCTTGCGCAGCT 167
```

RESULT 10
LOCUS   AA573470   282 bp   mRNA   EST   12-SEP-1997
DEFINITION   nm53f09.s1 NCI_CGAP_Br2 Homo sapiens cDNA clone IMAGE:1071977, mRNA
             sequence.
ACCESSION   AA573470
NID   g2347998
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
             Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
             Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 282)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Tel: (301) 496-1550
             Email: Robert_Strausberg@nih.gov
             Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
             Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             www-bio.llnl.gov/bbrp/image/image.html

             Insert Length: 328 Std Error: 0.00
             Seq primer: -40m13 fwd. ET from Amersham
             High quality sequence stop: 99.
FEATURES   Location/Qualifiers
     source   1..282
                     /organism="Homo sapiens"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from pooled bulk
                     breast tumor tissue, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector. This library is the normalized version of
                     NCI_CGAP_Br1.1. Library was constructed by Bento Soares
                     and M. Fatima Bonaldo."
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1071977"
                     /clone_lib="NCI_CGAP_Br2"
                     /sex="female, pooled"
                     /tissue_type="breast"
                     /lab_host="DH10B"
     mRNA   <1. >282
BASE COUNT   79 a   55 c   49 g   99 t
ORIGIN

Query Match   1.4%; Score 30; DB 9; Length 282;

Best Local Similarity 75.0%; Pred. No. 8.08e-08;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Db    4 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGACAAGTGAGAAAGCTTATTAA 63
          |||| ||||||| || ||| || |||||||||||||||||| || |||||||
Cp 1863 TTTTATTTTTTATTTATTATTTTGCTTGCATAGTTTGACAAGTCTGAAGACTCTATTAA 1804
```

RESULT 11
LOCUS   B19588   707 bp   DNA   GSS   15-SEP-1997
DEFINITION   T2203-T7.1 TAMU Arabidopsis thaliana genomic clone T2203, genomic
             survey sequence.
ACCESSION   B19588
NID   g2394642
KEYWORDS   GSS.
SOURCE   thale cress.
ORGANISM   Arabidopsis thaliana
             Eukaryotae; viridiplantae; Charophyta/Embryophyta group;
             Embryophyta; vascular plants; seed plants; Magnoliophyta;
             Magnoliopsida; Capparales; Brassicaceae; Arabidopsis.
REFERENCE   1 (bases 1 to 707)
AUTHORS   Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and
             Ecker,J.
TITLE   BAC End Sequences at ATGC
JOURNAL   Unpublished (1997)
COMMENT   Other_GSSs: T2203-T7, T2203-Sp6.1, T2203-Sp6, T2203-Sp6.2,
             T2203-Sp6.3, T2203-T7.2
             Contact: Ecker J.
             Arabidopsis Thaliana Genome Center
             University of Pennsylvania
             Dept. of Biology, University of Pennsylvania, Philadelphia, PA
             19104
             Tel: 215-898-9384
             Fax: 215-898-8780
             Email: jecker@atgenome.bio.upenn.edu
             Seq primer: T7
             Class: BAC ends
             High quality sequence start: 96
             High quality sequence stop: 181.
FEATURES   Location/Qualifiers
     source   1. .707
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /note="Vector: BeloBACII; Site_1: HindIII; Site_2:
                     HindIII; Produced by Rod Wing"
                     /db_xref="taxon:3702"
                     /clone="T2203"
                     /clone_lib="TAMU"
                     /sex="hermaphrodite"
BASE COUNT   136 a   149 c   78 g   208 t   136 others
ORIGIN

Best Local Similarity 1.4%; Score 31; DB 27; Length 707;
Matches 34; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

```
Db  485 TTTTATTTTTTTTATNTATTNTNTNNNNNNNNNNNNNNNNNNNNNNNNANNNA 544
          ||||||||| ||||| |||||
Cp 1863 TTTTATTTTTTATTTATTATTTTGCTTGCATAGTTTGACAAGTCTGAAGACTCTATTAA 1804

Db  545 NNNANAAANNANNANNNNNAACAANNAC 571
          ||| ||  ||  || ||
Cp 1803 GTTAAAAGAATAAAGAAATAAACTAC 1777
```

RESULT 12
LOCUS   AQ068616   189 bp   DNA   GSS   03-AUG-1998
DEFINITION   HS_2249_A1_H12_MF CIT Approved Human Genomic Sperm Library D Homo
             sapiens genomic clone Plate=2249 Col=23 Row=O, genomic survey
             sequence.
ACCESSION   AQ068616
NID   g3383815
KEYWORDS   GSS.

SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 189)
AUTHORS   Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
            Traicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
TITLE   Construction of a Characterized Clone Resource for Genomic
            Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
            Tagged Connectors
JOURNAL   Unpublished (1997)
COMMENT   Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2249 row: O column: 23
            Class: BAC ends
            High quality sequence stop: 189.
FEATURES   Location/Qualifiers
     source 1..189
            /organism="Homo sapiens"
            /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
            E-coli DH10B"
            /db_xref="taxon:9606"
            /clone="Plate=2249 Col=23 Row=O"
            /clone_lib="CIT Approved Human Genomic Sperm Library D"
            /sex="male"
BASE COUNT 40 a   48 c   30 g   69 t
ORIGIN

Query Match   1.3%; Score 28; DB 27; Length 189;
Best Local Similarity 96.78%; Pred. No. 1.16e-05;
Matches   29; Conservative   0; Mismatches   0; Indels   1; Gaps   2 others

Db   24 TATTCTTTTTATTTTATTATTATTTTG 53
          |||||||||||||||||||||||||
Cp   1869 TATTCATTTTATTTTATTTTATTATTATTTTG 1840

RESULT   13
LOCUS   AQ022111   407 bp   DNA   GSS   08-JUN-1998
DEFINITION CIT-HSP-2303D3.TR CIT-HSP Homo sapiens genomic clone 2303D3,
            genomic survey sequence.
ACCESSION   AQ022111
NID   g3200847
KEYWORDS   GSS.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 407)
AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
TITLE   Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
JOURNAL   Unpublished (1997)
COMMENT   Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.

Seq primer: M13 Reverse
Class: BAC ends.
FEATURES   Location/Qualifiers
     source 1..407
            /organism="Homo sapiens"
            /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
            HindIII"
            /db_xref="taxon:9606"
            /clone="2303D3"
            /clone_lib="CIT-HSP"
            /sex="Male"
            /cell_type="Sperm"
BASE COUNT 135 a   60 c   72 g   140 t
ORIGIN

Query Match   1.3%; Score 29; DB 27; Length 407;
Best Local Similarity 77.4%; Pred. No. 9.96e-07;
Matches   48; Conservative   0; Mismatches   13; Indels   1; Gaps   1;

Db   34 TGAGAGCTCTTACATATTTAGAAACAACTACTTTACCACATATATGATTTGCAGTACT 93
          |||||| |||| ||||| |||||||| ||| ||| | ||||||| ||||| || |||
Cp   1662 TGAGACATTTTAAATATTT-GAATCAAATACTATAGCACAAATATAATTTACAAAACT 1604

Db   94 TT 95
          ||
Cp   1603 TT 1602

RESULT   14
LOCUS   AQ046443   419 bp   DNA   GSS   13-JUL-1998
DEFINITION RPCI11-34J1.TK RPCI11 Homo sapiens genomic clone R-34J1, genomic
            survey sequence.
ACCESSION   AQ046443
NID   g3315370
KEYWORDS   GSS.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 419)
AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
            Venter,J.C.
TITLE   Use of BAC End Sequences for Sequence-Ready Map Building
JOURNAL   Unpublished (1997)
COMMENT   Other_GSSs: RPCI11-34J1.TJ
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            For clone availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Class: BAC ends.
FEATURES   Location/Qualifiers
     source 1..419
            /organism="Homo sapiens"
            /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
            RPCI11 Human Male BAC Library"
            /db_xref="taxon:9606"
            /clone="R-34J1"
            /clone_lib="RPCI11"
            /sex="Male"
            /cell_type="Lymphocytes"
BASE COUNT 131 a   73 c   76 g   139 t
ORIGIN

Query Match   1.3%; Score 29; DB 27; Length 419;
Best Local Similarity 77.4%; Pred. No. 9.96e-07;

Job time : 3820 secs.

```
Matches    48;   Conservative    0;   Mismatches   13;   Indels    1;   Gaps    1;

Db    343  ATATATGTGTTAC-ATATAAAGTGTAAATATATATTGTTGAGTGAGTTTTCTGTGCAT  401
           |||||||||| ||  |||||||| ||  |||| |||||| |  || ||||||||||||
Qy   1544  ATATATGTGTATTTTTATATAATGTCTAAAGTTATATTTCAGGTGTAATGTTTTCTGTGCAA  1603

Db    402  AG  403
           ||
Qy   1604  AG  1605
```

```
RESULT       15
LOCUS        AI041106        488 bp      mRNA              EST        30-JUN-1998
DEFINITION   ov61f03.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1641821
             3', mRNA sequence.
ACCESSION    AI041106
NID          g3280300
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
             Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 488)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Tel: (301) 496-1550
             Email: Robert_Strausberg@nih.gov
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
             Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             www-bio.llnl.gov/bbrp/image.html

             Seq primer: -40m13 fwd. ET from Amersham
             High quality sequence stop: 466.
FEATURES             Location/Qualifiers
     source          1..488
                     /organism="Homo sapiens"
                     /note="vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from Clontech
                     Laboratories, Inc., and primed with a Not I - oligo(dT)
                     primer [5'
                     TGTTACCCAATCTGAAGTGGGAGCGGCGGCCGCCAATTTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1641821"
                     /clone_lib="Soares_testis_NHT"
                     /sex="male"
                     /lab_host="DH10B"
BASE COUNT      161 a     82 c     76 g     169 t
ORIGIN
Query Match           1.3%;     Score 28;    DB 18;    Length 488;
Best Local Similarity 73.3%;    Pred. No. 1.16e-05;
Matches    44;   Conservative    0;   Mismatches   16;   Indels    0;   Gaps    0;

Db     22  TTGGTTTTTTTTTTTTGTAATAAACTTTGTAGAATAGGTTAACAACCAAAAGACAATACA  81
           || ||||  |||||||| ||||||| |||| |||||||||| ||||||||||||||| ||
Qy   2046  TTTGTTTTCTTTTTTTTTCCCATAGAAAAATGTACTATAGTTTATCAGCCAAAAAAACAATCCA  2105
```

Search completed: Sat Nov 14 03:12:39 1998

```
  _   _  ____  ____  ____   ____ _   _
 | \ | ||  _ \/ ___||  _ \ / ___| | | |
 |  \| || |_) \___ \| |_) | |   | |_| |
 | |\  ||  __/ ___) |  _ <| |___|  _  |
 |_| \_||_|   |____/|_| \_\\____|_| |_| (TM)
```

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sat Nov 14 03:21:01 1998; MasPar time 84.28 Seconds
                  1212.676 Million cell updates/sec
Tabular output not generated.

Title: >US-08-989-362-1
Description: (1-2191) from US08989362.seq
Perfect Score: 2191
N.A. Sequence: 2191 1 GCCAGGACCTCTGTGAACCG.........TTTTGGTACTTAAAAATGGC 2191
Comp: CGGTCCTGGACACACTTGGC.........AAAACCATGAATTTTTACCG

Scoring table: TABLE default
                  Gap 6

Nmatch STD :   Dbase 0; Query 0

Searched: 88822 seqs, 23323279 bases x 2

Post-processing: Minimum Match 0%
                  Listing first 45 summaries

Database: n-issued
                  1:5_COMB 2:PCT9_COMB 3:backfiles1

Statistics: Mean 9.409; Variance 5.535; scale 1.700

   Pred. No. is the number of results predicted by chance to have a
   score greater than or equal to the score of the result being printed,
   and is derived by analysis of the total score distribution.
```

SUMMARIES

| | Result No. | Score | % Match | Query Length | DB | ID | Description | Pred. No. |
|---|---|---|---|---|---|---|---|---|
| c | 1 | 47 | 2.1 | 7218 | 1 | US-08-232- | Sequence 14, Applicati | 1.14e-13 |
| c | 2 | 41 | 1.9 | 7218 | 1 | US-08-232- | Sequence 14, Applicati | 4.08e-10 |
| c | 3 | 39 | 1.8 | 215 | 1 | US-08-238- | Sequence 5, Applicatio | 5.86e-09 |
|   | 4 | 35 | 1.6 | 215 | 1 | US-08-238- | Sequence 5, Applicatio | 1.07e-06 |
|   | 5 | 26 | 1.2 | 85 | 2 | PCT-US94-0 | Sequence 26, Applicatio | 6.15e-02 |
|   | 6 | 26 | 1.2 | 85 | 2 | PCT-US94-0 | Sequence 1, Applicatio | 6.15e-02 |
|   | 7 | 26 | 1.2 | 85 | 1 | US-08-438- | Sequence 26, Applicatio | 6.15e-02 |
|   | 8 | 26 | 1.2 | 85 | 1 | US-08-370- | Sequence 1, Applicatio | 6.15e-02 |
|   | 9 | 27 | 1.2 | 105 | 1 | US-07-865- | Sequence 13, Applicati | 1.95e-02 |
|   | 10 | 24 | 1.1 | 66 | 1 | US-08-471- | Sequence 144, Applicat | 5.78e-01 |
|   | 11 | 24 | 1.1 | 69 | 1 | US-08-471- | Sequence 142, Applicat | 5.78e-01 |
|   | 12 | 24 | 1.1 | 74 | 2 | PCT-US95-1 | Sequence 94, Applicatio | 5.78e-01 |
|   | 13 | 24 | 1.1 | 74 | 2 | PCT-US95-1 | Sequence 100, Applicat | 5.78e-01 |
| c | 14 | 24 | 1.1 | 74 | 2 | PCT-US95-1 | Sequence 94, Applicatio | 5.78e-01 |
|   | 15 | 24 | 1.1 | 75 | 2 | PCT-US95-1 | Sequence 99, Applicatio | 5.78e-01 |
| c | 16 | 24 | 1.1 | 75 | 2 | PCT-US95-1 | Sequence 99, Applicatio | 5.78e-01 |
| c | 17 | 24 | 1.1 | 81 | 2 | PCT-US95-1 | Sequence 92, Applicatio | 5.78e-01 |
|   | 18 | 24 | 1.1 | 81 | 2 | PCT-US95-1 | Sequence 98, Applicati | 5.78e-01 |
| c | 19 | 25 | 1.1 | 81 | 2 | PCT-US95-1 | Sequence 92, Applicati | 1.91e-01 |
| c | 20 | 24 | 1.1 | 82 | 2 | PCT-US95-1 | Sequence 97, Applicati | 5.78e-01 |
| c | 21 | 24 | 1.1 | 82 | 2 | PCT-US95-1 | Sequence 97, Applicati | 5.78e-01 |
|   | 22 | 24 | 1.1 | 4131 | 2 | PCT-US92-0 | Sequence 3, Applicatio | 5.78e-01 |
|   | 23 | 25 | 1.1 | 7970 | 1 | US-08-135- | Sequence 31, Applicatio | 1.91e-01 |
| c | 24 | 23 | 1.0 | 65 | 1 | US-08-471- | Sequence 145, Applicat | 4.93e+00 |
|   | 25 | 23 | 1.0 | 66 | 1 | US-08-471- | Sequence 144, Applicat | 1.71e+00 |
|   | 26 | 23 | 1.0 | 68 | 1 | US-08-471- | Sequence 143, Applicat | 1.71e+00 |
| c | 27 | 23 | 1.0 | 105 | 1 | US-07-865- | Sequence 13, Applicatio | 1.71e+00 |
|   | 28 | 22 | 1.0 | 1312 | 2 | PCT-US94-0 | Sequence 1, Applicatio | 4.93e+00 |
|   | 29 | 22 | 1.0 | 1817 | 1 | US-08-473- | Sequence 5, Applicatio | 4.93e+00 |
|   | 30 | 22 | 1.0 | 2060 | 1 | PCT-US91-0 | Sequence 2, Applicatio | 4.93e+00 |
|   | 31 | 22 | 1.0 | 2107 | 1 | US-07-721- | Sequence 31, Applicati | 1.71e+00 |
|   | 32 | 23 | 1.0 | 2504 | 1 | US-08-390- | Sequence 1, Applicatio | 1.71e+00 |
|   | 33 | 23 | 1.0 | 2504 | 1 | US-08-484- | Sequence 15, Applicati | 1.71e+00 |
|   | 34 | 23 | 1.0 | 2823 | 1 | US-08-484- | Sequence 1, Applicatio | 1.71e+00 |
|   | 35 | 22 | 1.0 | 3024 | 1 | US-08-398- | Sequence 1, Applicatio | 4.93e+00 |
|   | 36 | 22 | 3 | | | 528493-1 | Patent No. 5284931. | 4.93e+00 |
|   | 37 | 23 | 1.0 | 3136 | 1 | US-08-332- | Sequence 41, Applicati | 1.71e+00 |
|   | 38 | 23 | 1.0 | 3136 | 1 | US-08-188- | Sequence 1, Applicatio | 1.71e+00 |
|   | 39 | 22 | 1.0 | 3291 | 2 | PCT-US94-0 | Sequence 41, Applicatio | 4.93e+00 |
|   | 40 | 23 | 1.0 | 3590 | 2 | PCT-US96-0 | Sequence 1, Applicatio | 1.71e+00 |
| c | 41 | 23 | 1.0 | 3590 | 2 | PCT-US96-0 | Sequence 1, Applicatio | 1.71e+00 |
|   | 42 | 22 | 1.0 | 4032 | 1 | US-08-107- | Sequence 3, Applicatio | 4.93e+00 |
| c | 43 | 22 | 1.0 | 4032 | 2 | PCT-US92-0 | Sequence 2, Applicatio | 4.93e+00 |
| c | 44 | 22 | 1.0 | 5852 | 1 | US-07-867- | Sequence 15, Applicati | 4.93e+00 |
| c | 45 | 22 | 1.0 | 17327 | 1 | US-07-906- | | 4.93e+00 |

ALIGNMENTS

```
RESULT 1
ID   US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
AC   xxxxxx
DT
DE   Sequence 14, Application US/08232463
CC   Sequence 14, Application US/08232463
CC   Patent No. 5670367
CC   GENERAL INFORMATION:
CC      APPLICANT: DORNER, F.
CC      APPLICANT: SCHEIFLINGER, F.
CC      APPLICANT: FALKNER, F.G.
CC      TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
CC      NUMBER OF SEQUENCES: 52
CC      CORRESPONDENCE ADDRESS:
CC         ADDRESSEE: Foley & Lardner
CC         STREET: 1800 Diagonal Road, Suite 500
CC         CITY: Alexandria
CC         STATE: VA
CC         COUNTRY: USA
CC         ZIP: 22313-0299
CC      COMPUTER READABLE FORM:
CC         MEDIUM TYPE: Floppy disk
CC         COMPUTER: IBM PC compatible
CC         OPERATING SYSTEM: PC-DOS/MS-DOS
CC         SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/08/232,463
CC         FILING DATE:
CC         CLASSIFICATION: 435
CC      PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER: US/07/935,313
CC         FILING DATE:
CC         APPLICATION NUMBER: EP 91 114 300.6
CC         FILING DATE: 26-AUG-1991
CC      ATTORNEY/AGENT INFORMATION:
CC         NAME: BENT, Stephen A.
CC         REGISTRATION NUMBER: 29,768
CC         REFERENCE/DOCKET NUMBER: 30472/114 IMMU
CC      TELECOMMUNICATION INFORMATION:
CC         TELEPHONE: (703)836-9300
CC         TELEFAX: (703)683-4109
CC         TELEX: 899149
CC      INFORMATION FOR SEQ ID NO: 14:
CC      SEQUENCE CHARACTERISTICS:
```

CC   LENGTH: 7218 base pairs
CC   TYPE: nucleic acid
CC   STRANDEDNESS: single
CC   TOPOLOGY: linear
CC   IMMEDIATE SOURCE:
CC   CLONE: pTz9pt-F1s
SQ   SEQUENCE 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.

Query Match   2.1%; Score 47; DB 1; Length 7218;
Best Local Similarity 7.9%; Pred. No. 1.14e-13;
Matches 14; Conservative 98; Mismatches 65; Indels 0; Gaps 0;

```
Db  1045 GCAGGTCGAGGAGCTGCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1104
         ||||| ||  | |:
Qy   162 GCAGCTCGGAGGAGATGGCAGCGGCCCCCGGCGTCCCACACGAGGGTCCGCTGCACCCCG 221
Db  1105 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1164
         :::::::::::::::::::::::::::::::::: :::: :
Qy   222 CGCCTTCTGCACCGGCTCCGGCGCCGCACCCGCGCCTCCCGCTCCATGTTCCTGGCCC 281
Db  1165 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1221
         :::       ::  :::
Qy   282 TCCTGGGGCTGGGACTGGCCCAGGTGGTCTGCAGCATCGCTCGTTCCTGTACTTTC 338
```

RESULT 2
ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
AC xxxxxx
DT
DE Sequence 14, Application US/08232463
CC Sequence 14, Application US/08232463
CC Patent No. 5670367
CC GENERAL INFORMATION:
CC   APPLICANT: DORNER, F.
CC   APPLICANT: SCHEIFLINGER, F.
CC   APPLICANT: FALKNER, F. G.
CC   TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
CC   NUMBER OF SEQUENCES: 52
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE: Foley & Lardner
CC   STREET: 1800 Diagonal Road, Suite 500
CC   CITY: Alexandria
CC   STATE: VA
CC   COUNTRY: USA
CC   ZIP: 22313-0299
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE: Floppy disk
CC   COMPUTER: IBM PC compatible
CC   OPERATING SYSTEM: PC-DOS/MS-DOS
CC   SOFTWARE: PatentIn Release #1.0, Version #1.25
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: US/08/232,463
CC   FILING DATE:
CC   CLASSIFICATION: 435
CC   PRIOR APPLICATION DATA:
CC   APPLICATION NUMBER: US/07/935,313
CC   FILING DATE:
CC   APPLICATION NUMBER: EP 91 114 300.6
CC   FILING DATE: 26-AUG-1991
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME: BENT, Stephen A.
CC   REGISTRATION NUMBER: 29,768
CC   REFERENCE/DOCKET NUMBER: 30472/114 IMMU
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE: (703)836-9300
CC   TELEFAX: (703)683-4109
CC   TELEX: 899149
CC INFORMATION FOR SEQ ID NO: 14:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH: 7218 base pairs
CC   TYPE: nucleic acid
CC   STRANDEDNESS: single
CC   TOPOLOGY: linear

CC   IMMEDIATE SOURCE:
CC   CLONE: pTz9pt-F1s
SQ   SEQUENCE 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.

Query Match   1.9%; Score 41; DB 1; Length 7218;
Best Local Similarity 0.5%; Pred. No. 4.08e-10;
Matches 1; Conservative 113; Mismatches 73; Indels 0; Gaps 0;

```
Db  1249 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1308
         ::: ::: ::: ::: :::::: :::::  :::: :: :::  :::: :
Cp  2012 CCTGTTCAATAATTGCATAAATACTCCTGACAGATTTCAGTAGCCTATTATGTATCCTT 1953
Db  1309 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1368
         ::: :::::::::::::::::::::::: :: ::: ::
Cp  1952 TTTGGCTATGTCAGCTCCTAAAGTCAATTAGCTTCTAAATTTGAAAGGCACCTGGTGACC 1893
Db  1369 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1428
         :::: :::::::: :::: ::::::::::::::::::::::::::::::
Cp  1892 AACATCCTACTTATTATTCAAGATATTCATTTTATTTTATTTATTTTGCTTGCAT 1833
Db  1429 YYYYYG 1435
         :::::|
Cp  1832 AGTTTTG 1826
```

RESULT 3
ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
AC xxxxxx
DT
DE Sequence 5, Application US/08238163
CC Sequence 5, Application US/08238163
CC Patent No. 5569830
CC GENERAL INFORMATION:
CC   APPLICANT: BENNETT, Alan
CC   APPLICANT: LABAVITCH, John M.
CC   APPLICANT: POWELL, Ann
CC   APPLICANT: STOTZ, Henrik
CC   TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
CC   TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEA
CC   NUMBER OF SEQUENCES: 24
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE: Townsend and Townsend Khourie and Crew
CC   STREET: Steuart Street Tower, One Market Plaza
CC   CITY: San Francisco
CC   STATE: California
CC   COUNTRY: US
CC   ZIP: 94105-1493
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE: Floppy disk
CC   COMPUTER: IBM PC compatible
CC   OPERATING SYSTEM: PC-DOS/MS-DOS
CC   SOFTWARE: PatentIn Release #1.0, Version #1.25
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: US/08/238,163
CC   FILING DATE: 03-MAY-1994
CC   CLASSIFICATION: 800
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME: Bastian, Kevin L.
CC   REGISTRATION NUMBER: 34,774
CC   REFERENCE/DOCKET NUMBER: 2307E-540
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE: (415) 543-9600
CC   TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH: 215 base pairs
CC   TYPE: nucleic acid
CC   STRANDEDNESS: single
CC   TOPOLOGY: unknown
CC   MOLECULE TYPE: protein
CC   FEATURE:
CC   NAME/KEY: misc_feature
CC   LOCATION: 1..215

```
CC   OTHER INFORMATION:   /standard_name= "Deduced amino acid
CC   OTHER INFORMATION:   sequence of PGIP from bean."
SQ   SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match            1.8%;      Score 39; DB 1; Length 215;
Best Local Similarity 12.5%;      Pred. No. 5.86e-09;
Matches  26; Conservative  85; Mismatches  96; Indels  1; Gaps  1;

Db    8  SSSVVSRTASCNDKAKKDGNTTSSWTDCCNRTWGVCDTDTTYRVNNDSGHNKYSSANYN  67
Cp  385  GCAGTGAGTGCGTCTTCTGATATTCTGTTAGGATCCATCTGCGCTCGAAAGTACAGGAA 326
Db   68  YGGNNVGAAKTHYTHTNVSGADSKTVTDSYNASGTSSSNGGTDGNRSGADSYGSSKTAM 127
Cp  325  CAGAGCGATGCTGCAGACCACCTGGCCCACTGCCCAGTCCCAGCCCCAGAGGGCCAGGAACATGGA 266
Db  128  TSRNRTGKTANNAVDSRNMGDASV-GSDKNTKKHAKNSADGKVGSKNNGDRNNRYGTGTK 186
Cp  265  GCGGGAGGCGGGGTTGGCGGGCGTGGGCGGAGCCCGGTGCAGAAGGCGCGGGGTGCAGCGGACC 206
Db  187  SNVSNNCGGGNKRDVSSYANNKCCGSSC 214
Cp  205  CTCGTGTGGGACGCCGGGCCCGTGCCC 178
```

```
RESULT 4
ID   US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
AC   xxxxxx
DT
DE   Sequence 5, Application US/08238163
CC   Sequence 5, Application US/08238163
CC   Patent No. 5569830
CC   GENERAL INFORMATION:
CC      APPLICANT: BENNETT, Alan
CC      APPLICANT: LABAVITCH, John M.
CC      APPLICANT: POWELL, Ann
CC      APPLICANT: STOTZ, Henrik
CC      TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
CC      TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
CC      NUMBER OF SEQUENCES: 24
CC      CORRESPONDENCE ADDRESS:
CC         ADDRESSEE: Townsend and Townsend Khourie and Crew
CC         STREET: Steuart Street Tower, One Market Plaza
CC         CITY: San Francisco
CC         STATE: California
CC         COUNTRY: US
CC         ZIP: 94105-1493
CC      COMPUTER READABLE FORM:
CC         MEDIUM TYPE: Floppy disk
CC         COMPUTER: IBM PC compatible
CC         OPERATING SYSTEM: PC-DOS/MS-DOS
CC         SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: US/08/238,163
CC         FILING DATE: 03-MAY-1994
CC         CLASSIFICATION: 800
CC      ATTORNEY/AGENT INFORMATION:
CC         NAME: Bastian, Kevin L.
CC         REGISTRATION NUMBER: 34,774
CC         REFERENCE/DOCKET NUMBER: 2307E-540
CC      TELECOMMUNICATION INFORMATION:
CC         TELEPHONE: (415) 543-9600
CC         TELEFAX: (415) 543-5043
CC   INFORMATION FOR SEQ ID NO: 5:
CC      SEQUENCE CHARACTERISTICS:
CC         LENGTH: 215 base pairs
CC         TYPE: nucleic acid
CC         STRANDEDNESS: single
CC         TOPOLOGY: unknown
CC      MOLECULE TYPE: protein
CC      FEATURE:
CC         NAME/KEY: misc_feature
CC         LOCATION: 1..215
```

```
CC   OTHER INFORMATION:   /standard_name= "Deduced amino acid
CC   OTHER INFORMATION:   sequence of PGIP from bean."
SQ   SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match            1.6%;      Score 35; DB 1; Length 215;
Best Local Similarity 15.9%;      Pred. No. 1.07e-06;
Matches  33; Conservative  83; Mismatches  87; Indels  4; Gaps  4;

Db   11  VVSRTASCNDKAKKDGNTTSSWTDCCNRTWGVCDTDTTYRVNNDSG-HNKYSSANYNYG  69
Qy   17  ACCGGTCGGGCGGGCGGGCCCCTGGCCGGGAGTCTCGCCGTGGTGGCCGAGGAAG  76
Db   70  GNNV-GAAKTHYTHTNVSGADSKTVTDSYNASGTSSSNGGTDGNRSGADSYGSSKTAMT  128
Qy   77  GGAGAACGATCGCGGAGCAGGCGCGCGCGACTCCCGGCGCGCGCGCCATGCGCCGGGC- 135
Db  129  SRNRTGKTANNAVDSRNMGDASVGSDKNTK-KHAKNSADGKYGSKNNGDRNNRYGTGTKS  187
Qy  136  CAGCCGAGACTACGGCAAGTACCTGCGCCAGCTCGGAGGAGATGGGCAGCGGCCCGGCGT 195
Db  188  NVSNNCGGGNKRDVSSYANNKCCGSSC 214
Qy  196  CCCACCACGAGGGTCCGCTGCACCCCGC 222
```

```
RESULT 5
ID   PCT-US94-05684-26 STANDARD; DNA; UNC; 85 BP.
AC   xxxxxx
DT
DE   Sequence 26, Application PC/TUS9405684
CC   Sequence 26, Application PC/TUS9405684
CC   GENERAL INFORMATION:
CC      APPLICANT: Chiron Corporation
CC      TITLE OF INVENTION: Peptide Inhibitors of Urokinase Receptor
CC      TITLE OF INVENTION: Activity
CC      NUMBER OF SEQUENCES: 37
CC      CORRESPONDENCE ADDRESS:
CC         ADDRESSEE: Chiron Corporation
CC         STREET: 4560 Horton Street
CC         CITY: Emeryville
CC         STATE: CA
CC         COUNTRY: USA
CC         ZIP: 94608
CC      COMPUTER READABLE FORM:
CC         MEDIUM TYPE: Floppy disk
CC         COMPUTER: IBM PC compatible
CC         OPERATING SYSTEM: PC-DOS/MS-DOS
CC         SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER: PCT/US94/05684
CC         FILING DATE: 19 MAY 1994
CC      CLASSIFICATION:
CC      ATTORNEY/AGENT INFORMATION:
CC         NAME: Green, Grant D.
CC         REGISTRATION NUMBER: 31,259
CC         REFERENCE/DOCKET NUMBER: 0941.100
CC      TELECOMMUNICATION INFORMATION:
CC         TELEPHONE: 510-601-2706
CC         TELEFAX: 510-655-3542
CC   INFORMATION FOR SEQ ID NO:  26:
CC      SEQUENCE CHARACTERISTICS:
CC         LENGTH: 85 base pairs
CC         TYPE: nucleic acid
CC         STRANDEDNESS: single
CC         TOPOLOGY: linear
CC      MOLECULE TYPE: cDNA
CC      HYPOTHETICAL: NO
SQ   SEQUENCE 85 BP; 6 A; 20 C; 3 G; 11 T; 45 OTHER.

Query Match            1.2%;      Score 26; DB 2; Length 85;
Best Local Similarity 20.7%;      Pred. No. 6.15e-02;
Matches  12; Conservative  15; Mismatches  31; Indels  0; Gaps  0;
```

```
Db  15 TCCGCTGAANNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSCCGC 72
        ||||||| |  .  :  . :  . :  . :  . :  . :  .:||||
Qy 208 TCCGCTGCACCCCGCCGCCTTCTGCACCGGCTCCGGCGCCGCCACCCGGCGCCTCCCGC 265
```

RESULT 6
ID   PCT-US94-05591-1 STANDARD; DNA; UNC; 85 BP.
AC   xxxxxx
DT
DE   Sequence 1, Application PC/TUS9405591
CC   Sequence 1, Application PC/TUS9405591
CC   GENERAL INFORMATION:
CC   APPLICANT: Doyle, Michael V.
CC   TITLE OF INVENTION: Method For Selection Of Biologically
CC   TITLE OF INVENTION: Active Peptide Sequences
CC   NUMBER OF SEQUENCES: 21
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE: Chiron Corporation
CC   STREET: 4560 Horton Street
CC   CITY: Emeryville
CC   STATE: CA
CC   COUNTRY: USA
CC   ZIP: 94608
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE: Floppy disk
CC   COMPUTER: IBM PC compatible
CC   OPERATING SYSTEM: PC-DOS/MS-DOS
CC   SOFTWARE: PatentIn Release #1.0, Version #1.25
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: PCT/US94/05591
CC   FILING DATE: 19-MAY-1994
CC   CLASSIFICATION:
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME: Green, Grant D.
CC   REGISTRATION NUMBER: 31,259
CC   REFERENCE/DOCKET NUMBER: 0407.100
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE: 510-601-2706
CC   TELEFAX: 510-655-3542
CC   INFORMATION FOR SEQ ID NO: 1:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH: 85 base pairs
CC   TYPE: nucleic acid
CC   STRANDEDNESS: double
CC   TOPOLOGY: linear
CC   MOLECULE TYPE: DNA (genomic)
CC   HYPOTHETICAL: NO
CC   FEATURE:
CC   NAME/KEY: misc_feature
CC   LOCATION: 24..68
CC   OTHER INFORMATION: /function= "Random base sequence
CC   OTHER INFORMATION: /label= Randomer
CC   OTHER INFORMATION:
SQ   SEQUENCE 85 BP; 6 A; 20 C; 3 G; 11 T; 45 OTHER.

Query Match   1.2%; Score 26; DB 2; Length 85;
Best Local Similarity 20.7%; Pred. No. 6.15e-02;
Matches 12; Conservative 15; Mismatches 31; Indels 0; Gaps 0;

```
Db  15 TCCGCTGAANNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSCCGC 72
        ||||||| |  .  :  . :  . :  . :  . :  . :  .:||||
Qy 208 TCCGCTGCACCCCGCCGCCTTCTGCACCGGCTCCGGCGCCGCCACCCGGCGCCTCCCGC 265
```

RESULT 7
ID   US-08-438-759-26 STANDARD; DNA; UNC; 85 BP.
AC   xxxxxx
DT
DE   Sequence 26, Application US/08438759
CC   Sequence 26, Application US/08438759
CC   Patent No. 5679782
CC   GENERAL INFORMATION:
CC   APPLICANT: Rosenberg, Steven
CC   APPLICANT: Doyle, Michael
CC   APPLICANT: Goodson, Robert
CC   TITLE OF INVENTION: Peptide Inhibitors of Urokinase Receptor
CC   TITLE OF INVENTION: Activity
CC   NUMBER OF SEQUENCES: 37
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE: Chiron Corporation
CC   STREET: 4560 Horton Street
CC   CITY: Emeryville
CC   STATE: CA
CC   COUNTRY: USA
CC   ZIP: 94608
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE: Floppy disk
CC   COMPUTER: IBM PC compatible
CC   OPERATING SYSTEM: PC-DOS/MS-DOS
CC   SOFTWARE: PatentIn Release #1.0, Version #1.25
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: US/08/438,759
CC   FILING DATE:
CC   CLASSIFICATION: 536
CC   PRIOR APPLICATION DATA:
CC   APPLICATION NUMBER: US/08/370,567
CC   FILING DATE:
CC   APPLICATION NUMBER: US/08/061,514
CC   FILING DATE:
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME: Green, Grant D.
CC   REGISTRATION NUMBER: 31,259
CC   REFERENCE/DOCKET NUMBER: 0941.001
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE: 510-601-2706
CC   TELEFAX: 510-655-3542
CC   INFORMATION FOR SEQ ID NO: 26:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH: 85 base pairs
CC   TYPE: nucleic acid
CC   STRANDEDNESS: single
CC   TOPOLOGY: linear
CC   MOLECULE TYPE: CDNA
CC   HYPOTHETICAL: NO
SQ   SEQUENCE 85 BP; 6 A; 20 C; 3 G; 11 T; 45 OTHER.

Query Match   1.2%; Score 26; DB 1; Length 85;
Best Local Similarity 20.7%; Pred. No. 6.15e-02;
Matches 12; Conservative 15; Mismatches 31; Indels 0; Gaps 0;

```
Db  15 TCCGCTGAANNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSCCGC 72
        ||||||| |  .  :  . :  . :  . :  . :  . :  .:||||
Qy 208 TCCGCTGCACCCCGCCGCCTTCTGCACCGGCTCCGGCGCCGCCACCCGGCGCCTCCCGC 265
```

RESULT 8
ID   US-08-370-567-26 STANDARD; DNA; UNC; 85 BP.
AC   xxxxxx
DT
DE   Sequence 26, Application US/08370567
CC   Sequence 26, Application US/08370567
CC   Patent No. 5656726
CC   GENERAL INFORMATION:
CC   APPLICANT: Rosenberg, Steven
CC   APPLICANT: Doyle, Michael
CC   APPLICANT: Goodson, Robert
CC   TITLE OF INVENTION: Peptide Inhibitors of Urokinase Receptor
CC   TITLE OF INVENTION: Activity
CC   NUMBER OF SEQUENCES: 37
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE: Chiron Corporation
CC   STREET: 4560 Horton Street
CC   CITY: Emeryville
CC   STATE: CA
CC   COUNTRY: USA
CC   ZIP: 94608

```
CC  COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.25
CC  CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/370,567
CC      FILING DATE:
CC      CLASSIFICATION: 530
CC  PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/061,514
CC      FILING DATE:
CC  ATTORNEY/AGENT INFORMATION:
CC      NAME: Green, Grant D.
CC      REGISTRATION NUMBER: 31,259
CC      REFERENCE/DOCKET NUMBER: 0941.001
CC  TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: 510-601-2706
CC      TELEFAX: 510-655-3542
CC  INFORMATION FOR SEQ ID NO: 26:
CC  SEQUENCE CHARACTERISTICS:
CC      LENGTH: 85 base pairs
CC      TYPE: nucleic acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: cDNA
CC      HYPOTHETICAL: NO
SQ  SEQUENCE 85 BP: 6 A; 20 C; 3 G; 11 T; 45 OTHER.

Query Match           1.2%;   Score 26;  DB 1; Length 85;
Best Local Similarity 20.7%;  Pred. No. 6.15e-02;
Matches  12;  Conservative 15;  Mismatches 31;  Indels  0;  Gaps  0;

Db   15 TCCGCTGAANNSNNSNNSNNSNNSNNSNNSNNSNNSNNSCCGC 72
         ||||||     :  :        :  :       ::  : ||||
Qy  208 TCCGCTGCACCCCGGCGCTTCGCACGGCTCCGGCGCCACCCGGCGCCCTCCCGC 265


RESULT 9
ID  US-07-865-662F-13 STANDARD; DNA; UNC; 105 BP.
AC  xxxxxx
DT
DE  Sequence 13, Application US/07865662F
DE  Sequence 13, Application US/07865662F
CC  Patent No. 5451670
CC  GENERAL INFORMATION:
CC      APPLICANT: Marcia M. Miller
CC      TITLE OF INVENTION: Restriction Fragment Length
CC      TITLE OF INVENTION: Polymorphism Test For Haplotyping Domesticated Fowl
CC      NUMBER OF SEQUENCES: 14
CC  CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: City of Hope
CC      STREET: 1500 East Duarte Road
CC      CITY: Duarte
CC      STATE: California
CC      COUNTRY: United States of America
CC      ZIP: 91010-0269
CC  COMPUTER READABLE FORM:
CC      MEDIUM TYPE: 3M Double Density 5 1/4" diskette
CC      COMPUTER: Wang PC
CC      OPERATING SYSTEM: MS DOS Version 3.20
CC      SOFTWARE: Microsoft
CC  CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/07/865,662F
CC      FILING DATE: 07 April, 1992
CC      CLASSIFICATION: 435
CC  PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: 07/688,326
CC      FILING DATE: 22 April 1991
CC      APPLICATION NUMBER: 07/588,922
CC      FILING DATE: 27 September 1990
CC      APPLICATION NUMBER: 07/210,405
CC      FILING DATE: 23 June 1988


CC      APPLICATION NUMBER: US 07/130,529
CC      FILING DATE: 9 December 1987
CC      APPLICATION NUMBER: US 07/068,176
CC      FILING DATE: 30 June 1987
CC  ATTORNEY/AGENT INFORMATION:
CC      NAME: Irons, Edward S.
CC      REGISTRATION NUMBER: 16,541
CC      REFERENCE/DOCKET NUMBER: NO. 5451670e
CC  TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (202) 783-6040
CC      TELEFAX: (202) 783-6031
CC      TELEX: NO. 5451670e
CC  INFORMATION FOR SEQ ID NO: 13:
CC  SEQUENCE CHARACTERISTICS:
CC      LENGTH: 105
CC      TYPE: Nucleic Acid
CC      STRANDEDNESS: Double
CC      TOPOLOGY: Linear
CC      MOLECULE TYPE: DNA
CC      ORIGINAL SOURCE: Synthetically Prepared
CC      IMMEDIATE SOURCE: Synthetically Prepared
SQ  SEQUENCE 105 BP: 15 A; 0 C; 8 G; 1 T; 81 OTHER.

Query Match           1.2%;    Score 27;  DB 1; Length 105;
Best Local Similarity 10.4%;   Pred. No. 1.95e-02;
Matches  10;  Conservative 32;  Mismatches 54;  Indels  0;  Gaps  0;

Db    5 VGNNAKNSNNNKSNNAVKNGNNAKNSGKNKSANAVNNGNNAKNSNNMGTRDNKNN 64
          :| |: :  ::      ::  :::| | : : :|| : : :|| : :|| : :||
Qy   43 CGGGAGTCTGCTCGGCGGTGGTGGCCGGAGGAAGGGAGAGAGAACGATCGCGGAGCCGGAGCAGGGCG 102
Db   65 RNAAKNNNSGVADNKNNASNNYDNGSGVADNKNAAK 100
          : :  :|:   :::: :::: : :|  |:
Qy  103 CCCGAACTCCGGCGCCGGCGCCGCCATGCGCCGGCCAG 138


RESULT 10
ID  US-08-471-052A-144 STANDARD; DNA; UNC; 66 BP.
AC  xxxxxx
DT
DE  Sequence 144, Application US/08471052A
DE  Sequence 144, Application US/08471052A
CC  Patent No. 5625033
CC  GENERAL INFORMATION:
CC      APPLICANT: Kay, B. K.
CC      APPLICANT: Fowlkes, D. M.
CC      TITLE OF INVENTION: Totally Synthetic Affinity Reagents
CC      NUMBER OF SEQUENCES: 166
CC  CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Pennie & Edmonds
CC      STREET: 1155 Avenue of the Americas
CC      CITY: New York
CC      STATE: New York
CC      COUNTRY: U.S.A.
CC      ZIP: 10036-2711
CC  COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.25
CC  CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/471,052A
CC      FILING DATE: 06-JUNE-1995
CC      CLASSIFICATION: 530
CC  ATTORNEY/AGENT INFORMATION:
CC      NAME: Misrock, S. Leslie
CC      REGISTRATION NUMBER: 18,872
CC      REFERENCE/DOCKET NUMBER: 1101-179
CC  TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: 212 790-9090
CC      TELEFAX: 212 869-8864/9741
CC      TELEX: 66141 PENNIE
CC  INFORMATION FOR SEQ ID NO: 144:
```

```
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH:  66 bases
CC        TYPE:  nucleic acid
CC        STRANDEDNESS:  single
CC        TOPOLOGY:  unknown
CC      MOLECULE TYPE:  DNA
SQ      SEQUENCE 66 BP; 2 A; 3 C; 4 G; 2 T; 55 OTHER.

Query Match          1.1%;   Score 24;  DB 1;  Length 66;
Best Local Similarity   9.8%;   Pred. No. 5.78e-01;
Matches     6;  Conservative 18;  Mismatches 37;  Indels   0;  Gaps   0;

Db          2 CGAGNNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBC 61
              ||||     : : : :  : : : :  :  : :  :  : : :|
Qy        202 CGAGGGTCCGCTGCACCCGCGCCTTCTGCACCGGCTCCGGCGGCCGCCACCCGCCGCCTC 261

Db         62 C 62
              |
Qy        262 C 262

RESULT 11
ID      US-08-471-052A-142 STANDARD; DNA; UNC; 69 BP.
AC      xxxxxx
DT
DE      Sequence 142, Application US/08471052A
CC      Sequence 142, Application US/08471052A
CC      Patent No. 5625033
CC      GENERAL INFORMATION:
CC        APPLICANT:  Kay, B. K.
CC        APPLICANT:  Fowlkes, D. M.
CC        TITLE OF INVENTION:  Totally Synthetic Affinity Reagents
CC        NUMBER OF SEQUENCES:  166
CC        CORRESPONDENCE ADDRESS:
CC          ADDRESSEE:  Pennie & Edmonds
CC          STREET:  1155 Avenue of the Americas
CC          CITY:  New York
CC          STATE:  New York
CC          COUNTRY:  U.S.A.
CC          ZIP:  10036-2711
CC        COMPUTER READABLE FORM:
CC          MEDIUM TYPE:  Floppy disk
CC          COMPUTER:  IBM PC compatible
CC          OPERATING SYSTEM:  PC-DOS/MS-DOS
CC          SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  US/08/471,052A
CC          FILING DATE:  06-JUNE-1995
CC          CLASSIFICATION:  530
CC        ATTORNEY/AGENT INFORMATION:
CC          NAME:  Misrock, S. Leslie
CC          REGISTRATION NUMBER:  18,872
CC          REFERENCE/DOCKET NUMBER:  1101-179
CC        TELECOMMUNICATION INFORMATION:
CC          TELEPHONE:  212 790-9090
CC          TELEFAX:  212 869-8864/9741
CC          TELEX:  66141 PENNIE
CC      INFORMATION FOR SEQ ID NO:  142:
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH:  69 bases
CC          TYPE:  nucleic acid
CC          STRANDEDNESS:  single
CC          TOPOLOGY:  unknown
CC        MOLECULE TYPE:  DNA
SQ      SEQUENCE 69 BP; 2 A; 4 C; 6 G; 2 T; 55 OTHER.

Query Match          1.1%;   Score 24;  DB 1;  Length 69;
Best Local Similarity   9.8%;   Pred. No. 5.78e-01;
Matches     6;  Conservative 18;  Mismatches 37;  Indels   0;  Gaps   0;

Db          5 CGAGNNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBC 64
              ||||     : : : :  : : : :  :  : :  :  : : :|
Qy        202 CGAGGGTCCGCTGCACCCGCGCCTTCTGCACCGGCTCCGGCGGCCGCCACCCGCCGCCTC 261
```

```
Db         65 C 65
              |
Qy        262 C 262

RESULT 12
ID      PCT-US95-11934-94 STANDARD; DNA; UNC; 74 BP.
AC      xxxxxx
DT
DE      Sequence 94, Application PC/TUS9511934
CC      Sequence 94, Application PC/TUS9511934
CC      GENERAL INFORMATION:
CC        APPLICANT:  Cytogen Corporation
CC        TITLE OF INVENTION:  Antigen Binding Peptides (Abtides) From
CC        TITLE OF INVENTION:  Peptide Libraries
CC        NUMBER OF SEQUENCES:  103
CC        CORRESPONDENCE ADDRESS:
CC          ADDRESSEE:  Pennie & Edmonds
CC          STREET:  1155 Avenue of the Americas
CC          CITY:  New York
CC          STATE:  New York
CC          COUNTRY:  USA
CC          ZIP:  10036
CC        COMPUTER READABLE FORM:
CC          MEDIUM TYPE:  Floppy disk
CC          COMPUTER:  IBM PC compatible
CC          OPERATING SYSTEM:  PC-DOS/MS-DOS
CC          SOFTWARE:  PatentIn Release #1.0, Version #1.30
CC        CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  PCT/US95/11934
CC          FILING DATE:  20-SEP-1995
CC          CLASSIFICATION:
CC        ATTORNEY/AGENT INFORMATION:
CC          NAME:  Misrock, S. Leslie
CC          REGISTRATION NUMBER:  18,872
CC          REFERENCE/DOCKET NUMBER:  1101-196-228
CC        TELECOMMUNICATION INFORMATION:
CC          TELEPHONE:  (212) 790-9090
CC          TELEFAX:  (212) 869-9741/8864
CC          TELEX:  66141 PENNIE
CC      INFORMATION FOR SEQ ID NO:  94:
CC        SEQUENCE CHARACTERISTICS:
CC          LENGTH:  74 base pairs
CC          TYPE:  nucleic acid
CC          STRANDEDNESS:  single
CC          TOPOLOGY:  linear
CC        MOLECULE TYPE:  DNA (genomic)
SQ      SEQUENCE 74 BP; 3 A; 4 C; 3 G; 1 T; 63 OTHER.

Query Match          1.1%;   Score 24;  DB 2;  Length 74;
Best Local Similarity   6.2%;   Pred. No. 5.78e-01;
Matches     4;  Conservative 20;  Mismatches 41;  Indels   0;  Gaps   0;

Db          2 CGAGNNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBN 61
              ||||
Qy        202 CGAGGGTCCGCTGCACCCGCGCCTTCTGCACCGGCTCCGGCGGCCGCCACCCGCCGCCTC 261

Db         62 NBNNB 66
              :
Qy        262 CCGCT 266

RESULT 13
ID      PCT-US95-11934-100 STANDARD; DNA; UNC; 74 BP.
AC      xxxxxx
DT
DE      Sequence 100, Application PC/TUS9511934
CC      Sequence 100, Application PC/TUS9511934
CC      GENERAL INFORMATION:
CC        APPLICANT:  Cytogen Corporation
CC        TITLE OF INVENTION:  Antigen Binding Peptides (Abtides) From
CC        TITLE OF INVENTION:  Peptide Libraries
```

CC   NUMBER OF SEQUENCES: 103
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE: Pennie & Edmonds
CC   STREET: 1155 Avenue of the Americas
CC   CITY: New York
CC   STATE: New York
CC   COUNTRY: USA
CC   ZIP: 10036
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE: Floppy disk
CC   COMPUTER: IBM PC compatible
CC   OPERATING SYSTEM: PC-DOS/MS-DOS
CC   SOFTWARE: PatentIn Release #1.0, Version #1.30
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: PCT/US95/11934
CC   FILING DATE: 20-SEP-1995
CC   CLASSIFICATION:
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME: Misrock, S. Leslie
CC   REGISTRATION NUMBER: 18,872
CC   REFERENCE/DOCKET NUMBER: 1101-196-228
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE: (212) 790-9090
CC   TELEFAX: (212) 869-9741/8864
CC   TELEX: 66141 PENNIE
CC   INFORMATION FOR SEQ ID NO: 100:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH: 74 base pairs
CC   TYPE: nucleic acid
CC   STRANDEDNESS: single
CC   TOPOLOGY: linear
CC   MOLECULE TYPE: DNA (genomic)
SQ   SEQUENCE 74 BP: 6 A; 6 C; 1 G; 1 T; 60 OTHER.

Query Match   1.1%;   Score 24; DB 2; Length 74;
Best Local Similarity 10.3%; Pred. No. 5.78e-01;
Matches   7; Conservative   19; Mismatches 42; Indels   0; Gaps   0;

Db   3 AGAVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNN 62
        |||: : : : : : : : : : : : : : : : : : :
Qy 81 AGAACGATCCGGGAGCAGGGCGCCCAACTCCGGGGCGCCGCCCATGCGCGGGCCCAGCC 140

Db 63 VNNACCAC 70
           : ||||
Qy 141 GAGACTAC 148

RESULT 14
PCT-US95-11934-94 STANDARD; DNA; UNC; 74 BP.
xxxxxx
DE Sequence 94, Application PC/TUS9511934
DE Sequence 94, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC   APPLICANT: Cytogen Corporation
CC   TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC   TITLE OF INVENTION: Peptide Libraries
CC   NUMBER OF SEQUENCES: 103
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE: Pennie & Edmonds
CC   STREET: 1155 Avenue of the Americas
CC   CITY: New York
CC   STATE: New York
CC   COUNTRY: USA
CC   ZIP: 10036
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE: Floppy disk
CC   COMPUTER: IBM PC compatible
CC   OPERATING SYSTEM: PC-DOS/MS-DOS
CC   SOFTWARE: PatentIn Release #1.0, Version #1.30
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: PCT/US95/11934
CC   FILING DATE: 20-SEP-1995

CC   CLASSIFICATION:
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME: Misrock, S. Leslie
CC   REGISTRATION NUMBER: 18,872
CC   REFERENCE/DOCKET NUMBER: 1101-196-228
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE: (212) 790-9090
CC   TELEFAX: (212) 869-9741/8864
CC   TELEX: 66141 PENNIE
CC   INFORMATION FOR SEQ ID NO: 94:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH: 74 base pairs
CC   TYPE: nucleic acid
CC   STRANDEDNESS: single
CC   TOPOLOGY: linear
CC   MOLECULE TYPE: DNA (genomic)
SQ   SEQUENCE 74 BP; 3 A; 4 C; 3 G; 1 T; 63 OTHER.

Query Match   1.1%;   Score 24; DB 2; Length 74;
Best Local Similarity 7.4%; Pred. No. 5.78e-01;
Matches   5; Conservative   20; Mismatches 43; Indels   0; Gaps   0;

Db   4 AGNNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNB 63
        ||: :
Cp 324 AGAGCGATGCTGCAGACCACCTGGCCCAGTCCAGCCCCAGGAGGGCCCAGGAACATGGAG 265

Db 64 NNBNACGC 71
           : ||
Cp 264 CGGGAGGC 257

RESULT 15
ID PCT-US95-11934-99 STANDARD; DNA; UNC; 75 BP.
AC xxxxxx
DT
DE Sequence 99, Application PC/TUS9511934
DE Sequence 99, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC   APPLICANT: Cytogen Corporation
CC   TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC   TITLE OF INVENTION: Peptide Libraries
CC   NUMBER OF SEQUENCES: 103
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE: Pennie & Edmonds
CC   STREET: 1155 Avenue of the Americas
CC   CITY: New York
CC   STATE: New York
CC   COUNTRY: USA
CC   ZIP: 10036
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE: Floppy disk
CC   COMPUTER: IBM PC compatible
CC   OPERATING SYSTEM: PC-DOS/MS-DOS
CC   SOFTWARE: PatentIn Release #1.0, Version #1.30
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: PCT/US95/11934
CC   FILING DATE: 20-SEP-1995
CC   CLASSIFICATION:
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME: Misrock, S. Leslie
CC   REGISTRATION NUMBER: 18,872
CC   REFERENCE/DOCKET NUMBER: 1101-196-228
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE: (212) 790-9090
CC   TELEFAX: (212) 869-9741/8864
CC   TELEX: 66141 PENNIE
CC   INFORMATION FOR SEQ ID NO: 99:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH: 75 base pairs
CC   TYPE: nucleic acid
CC   STRANDEDNESS: single
CC   TOPOLOGY: linear
CC   MOLECULE TYPE: DNA (genomic)

SQ   SEQUENCE 75 BP; 1 A; 1 C; 7 G; 5 T; 61 OTHER.

Query Match   1.1%;   Score 24; DB 2; Length 75;
Best Local Similarity   6.2%;   Pred. No. 5.78e-01;
Matches   4; Conservative   20; Mismatches 41; Indels   0; Gaps   0;

Db   2 CGAGNNNBNBNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBN 61
          |||| : : : : : : : : : : : : : : : : :
Qy   202 CGAGGGTCCGCTGCACCCCGGCCTTCTGCACCGGCTCCGGCGCCGCCACCCCGCCGCCTC 261

Db   62 NBNNB 66
           : :
Qy   262 CCGCT 266

Search completed: Sat Nov 14 03:23:58 1998
Job time : 177 secs.

```
****************************************
*                                      *
*   MPSRCH   (TM)                      *
*                                      *
****************************************

         Release 3.1A John F. Collins, Biocomputing Research Unit.
             Copyright (C) 1993-1998 University of Edinburgh, U.K.
                   Distribution rights by Oxford Molecular Ltd

MPsrch_nn   n.a. - n.a. database search, using Smith-Waterman algorithm

Run on:           Sat Nov 14 03:12:59 1998;   MasPar time 320.41 Seconds
                                               930.348 Million cell updates/sec
Tabular output not generated.

Title:            >US-08-989-362-1
Description:      (1-2191) from US08989362.seq
Perfect Score:    2191
N.A. Sequence:    2191    1 GCCAGGACCTCTGTGAACCG..........TTTTGGTACTTAAAAATGGC 2191
                            CGGTCCTGGAGACACTTGGC..........AAAACCATGAATTTTTACCG
Comp:

Scoring table:    TABLE default
                  Gap 6

Nmatch   STD :

Dbase 0:   Query 0

Searched:         188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database:        n-geneseq32
                 1:part1  2:part2  3:part3  4:part4  5:part5  6:part6  7:part7
                 8:part8  9:part9 10:part10 11:part11 12:part12 13:part13
                14:part14 15:part15 16:part16 17:part17 18:part18
                19:part19 20:part20 21:part21 22:part22 23:part23
                24:part24 25:part25 26:part26 27:part27 28:part28
                29:part29 30:part30 31:part31 32:part32 33:part33
                34:part34 35:part35 36:part36 37:part37 38:part38
                39:part39 40:part40

Statistics:      Mean 9.870;  Variance 6.465;  scale 1.527

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                           SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	43	2.0	91	9	Q51746	Oligonucleotide probe	5.16e-09
2	41	1.9	204	1	N81164	Base substituted E.co	5.67e-08
3	40	1.8	91	9	Q51746	Oligonucleotide probe	1.86e-07
4	39	1.8	204	1	N81164	Base substituted E.co	6.03e-07
c 5	37	1.7	178	32	T76405	Human endothelin-1 an	6.19e-06
c 6	37	1.7	178	32	T76405	Human endothelin-1 an	6.19e-06
c 7	35	1.6	88	32	T76170	Human IL3 receptor an	6.11e-05
8	36	1.6	114	12	Q70469	Generic DNA sequence	1.96e-05
9	36	1.6	114	12	Q70465	Generic DNA sequence	1.96e-05
c 10	34	1.6	114	12	Q70467	Generic DNA sequence	1.89e-04
11	34	1.6	114	12	Q70468	Generic DNA sequence	1.89e-04
c 12	35	1.6	114	12	Q70467	Generic DNA sequence	6.11e-05
c 13	35	1.6	114	12	Q70469	Generic DNA sequence	6.11e-05

								Description	Pred. No.
c	14	34	1.6	114	12	Q70465	Generic DNA sequence	1.89e-04
c	15	34	1.6	114	12	Q70468	Generic DNA sequence	1.89e-04
	16	32	1.5	114	12	Q70466	Generic DNA sequence	1.74e-03
	17	32	1.5	114	12	Q70472	Generic DNA sequence	1.74e-03
c	18	32	1.5	114	12	Q70466	Generic DNA sequence	1.74e-03
	19	32	1.5	172	32	T76363	Human interleukin 8 a	1.74e-03
c	20	33	1.5	172	32	T76363	Human interleukin 8 a	5.78e-04
	21	33	1.5	190	32	T76452	Chymase antisense oli	5.78e-04
c	22	33	1.5	190	32	T76452	Chymase antisense oli	5.78e-04
c	23	31	1.5	317	32	T76274	Human neutrophil elas	5.78e-04
	24	31	1.4	100	32	T76186	Human IL4 receptor an	5.20e-03
	25	31	1.4	114	12	Q70471	Generic DNA sequence	5.20e-03
c	26	30	1.4	114	12	Q70470	Generic DNA sequence	5.20e-03
	27	30	1.4	114	12	Q70473	Generic DNA sequence	5.20e-03
c	28	31	1.4	114	12	Q70470	Generic DNA sequence	1.53e-03
c	29	30	1.4	128	32	T76233	Human IL6 antisense o	5.20e-03
c	30	31	1.4	162	32	T76307	Human RANTES antisens	1.53e-03
c	31	30	1.4	168	32	T76270	Human MDNCF antisense	5.20e-03
	32	30	1.4	201	2	N70195	Streptomyces protease	5.20e-03
c	33	30	1.4	498	3	N50027	Sequence encoding new	5.20e-03
c	34	31	1.4	501	3	N50032	Sequence encoding new	1.53e-03
c	35	30	1.4	501	3	N50031	Sequence encoding new	5.20e-03
c	36	30	1.4	501	3	N50028	Sequence encoding new	5.20e-03
c	37	30	1.4	501	3	N50033	Sequence encoding new	5.20e-03
c	38	30	1.4	501	3	N50029	Sequence encoding new	5.20e-03
c	39	30	1.4	501	3	N50029	Sequence encoding new	5.20e-03
c	40	28	1.3	114	12	Q70471	Generic DNA sequence	1.26e-01
	41	29	1.3	147	32	T76294	Human defensin 1 anti	4.43e-02
c	42	28	1.3	501	3	N50023	Sequence encoding new	1.26e-01
c	43	29	1.3	2368	1	Q04444	Sequence of wunl-85 g	4.43e-02
c	44	29	1.3	2368	4	Q23585	Anther specific promo	4.43e-02
c	45	29	1.3	2368	2	Q13372	Potato wunl gene with	4.43e-02

```
                           ALIGNMENTS

RESULT  1
ID   Q51746 standard; cDNA; 91 BP.
AC   Q51746;
DT   31-MAY-1994 (first entry)
DE   Oligonucleotide probe MK14-A
KW   Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW   ss.
OS   Synthetic.
PN   EP-571911-A.
PD   01-DEC-1993.
PF   24-MAY-1993; 108325.
PR   26-MAY-1992; US-889651.
PA   (BECT ) BECTON DICKINSON CO.
PI   Shank DD, Spears PA;
DR   WPI; 93-378844/48.
PT   New oligo:nucleotide probes specific for Mycobacteria - used for
PT   detection and amplification of Mycobacteria nucleic acid in
PT   samples
PS   Claim 3; Page 14; 23pp; English.
CC   Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC   (Q51735). It hybridized to all spp. of mycobacterial spp. tested, but
CC   cross reacted to a few non-mycobacterial spp. The probe may
CC   be useful as an initial screen for mycobacterial infection.
CC   See also Q51735-45 and Q51747-59.
SQ   Sequence   91 BP;   5 A;   17 C;   15 G;   4 T;

         Query Match     2.0%;
         Best Local Similarity  16.4%;
         Matches  10;  Conservative  42;  Mismatches  9;  Indels  0;  Gaps  0;

Db         1 ggctccggcggssvhsyyvhvvshhhsvhhvhhvvvhhvyhyhvhyyvsv  60
             |||||||||| :.:...::::::::::::::::  :.  ::::: .:::..
Qy       185 GGCCCCGGCGTCCCAACGAGAGGTCCCGTGCACCCCGGCGCCTTCGCACCGGCTCCGGCG 244

Db        61 c  61
             |
Qy       245 C 245
```

RESULT 2
ID   N81164 standard; DNA; 204 BP.
AC   N81164;
DT   08-NOV-1990 (first entry)
DE   Base substituted E.coli beta-galactosidase alpha-fragment.
KW   E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS   Escherichia coli.
FH   Key   Location/Qualifiers
FT   misc_feature   19..69
FT   /*tag= a
FT   /function=multiple cloning site
FT   primer_bind   187..204
FT   /*tag= b
PN   EP-285123-A.
PD   05-MAY-1988.
PF   30-MAR-1988; 105163.
PR   03-APR-1987; US-034819.
PA   (SUSO) SUOMEN SOKERI OY.
PI   Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
DR   WPI: 88-279927/40.
PT   Introducing random point mutations into nucleic acds -
PT   by prepn of single stranded template, annealing a primer, elongation,
PT   misincorporation, completion of molecules and screening.
PS   Disclosure; P; English.
CC   Random point mutations were introduced into the alpha fragment of
CC   E.coli beta-galactosidase. The wild type sequence was obtained as a
CC   single stranded template and an oligonucleotide was hybridised to
CC   it to generate a popn of DNA molecules which terminate at all
CC   possible nucleotide positions within a specified region. The
CC   variable 3' ends generated in this way are used as primers for
CC   reverse transcriptase. Nucleotides are misincorporated by the
CC   transcriptase and the molecules are completed to forms that can be
CC   amplified and then expressed in a suitable host-vector system.
CC   The sequence covers all 176 difft base substitutions, most of which
CC   occurred singularly in any given mutant.
CC   See also P80575.
SQ   Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match   1.9%; Score 41; DB 1; Length 204;
Best Local Similarity 11.7%; Pred. No. 5.67e-08;
Matches   12; Conservative   53; Mismatches   38; Indels   0; Gaps   0;

Db   80 hvgccgymrttthyrrmrbnvyrdynrsdaaawyccyrrsvkydccynachdddhyvyb 139
               ::|| :::|:|::|::|| || :|:|::: ::||| :: :::|: | :||:::|:
Cp   776 TGGCGTACAGGTAATAGAAGCCATCTTGGTTAACCCTTAGTTTCCGTTGCTTAACGTCA 717

Db   140 bbvynvhnhnncccbnnhvchnvhbnnhrnwayvrhdarrd 182
               ||::::::::::|: :|: :|: ::|||::::::::||::
Cp   716 TGTTAGAGATCTTGGCCCAGCCTCGATCGTGGTACCAAGAGGA 674

RESULT 3
ID   Q51746 standard; cDNA; 91 BP.
AC   Q51746;
DT   31-MAY-1994 (first entry)
DE   Oligonucleotide probe MK14-A
KW   Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW   ss.
OS   Synthetic.
PN   EP-571911-A.
PD   01-DEC-1993.
PF   24-MAY-1993; 108325.
PR   26-MAY-1992; US-889651.
PA   (BECT) BECTON DICKINSON CO.
PI   Shank DD, Spears PA;
DR   WPI: 93-378844/48.
PT   New oligo:nucleotide probes specific for Mycobacteria - used for
PT   detection and amplification of Mycobacteria nucleic acid in
PT   samples
PT   Claim 3; Page 14; 23pp; English.
CC   Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14

CC   (Q51735). It hybridized to all spp. of mycobacteria tested, but
CC   cross reacted to a few non-mycobacterial spp. The probe may
CC   be useful as an initial screen for mycobacterial infection.
CC   See also Q51735-45 and Q51747-59.
SQ   Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match   1.8%; Score 40; DB 9; Length 91;
Best Local Similarity   5.6%; Pred. No. 1.86e-07;
Matches   3; Conservative   44; Mismatches   7; Indels   0; Gaps   0;

Db   11 ssvhsyyvhvvshhhsvhhvhhvhsvvvhhvhvvhvhhvhyhvysvctca 64
               ::| :::||:|::::|: :|:::|: :|:::::||::::::::|::: |||
Cp   1509 GCAGGTCCAGGCAATGTAACATTCAAAAGAACTTAACCCTTCACACTATCA 1456

RESULT 4
ID   N81164 standard; DNA; 204 BP.
AC   N81164;
DT   08-NOV-1990 (first entry)
DE   Base substituted E.coli beta-galactosidase alpha-fragment.
KW   E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS   Escherichia coli.
FH   Key   Location/Qualifiers
FT   misc_feature   19..69
FT   /*tag= a
FT   /function=multiple cloning site
FT   primer_bind   187..204
FT   /*tag= b
PN   EP-285123-A.
PD   05-MAY-1988.
PF   30-MAR-1988; 105163.
PR   03-APR-1987; US-034819.
PA   (SUSO) SUOMEN SOKERI OY.
PI   Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
DR   WPI: 88-279927/40.
PT   Introducing random point mutations into nucleic acods -
PT   by prepn of single stranded template, annealing a primer, elongation,
PT   misincorporation, completion of molecules and screening.
PS   Disclosure: p; English.
CC   Random point mutations were introduced into the alpha fragment of
CC   E.coli beta-galactosidase. The wild type sequence was obtained as a
CC   single stranded template and an oligonucleotide was hybridised to
CC   it to generate a popn of DNA molecules which terminate at all
CC   possible nucleotide positions within a specified region. The
CC   variable 3' ends are generated in this way are used as primers for
CC   reverse transcriptase. Nucleotides are misincorporated by the
CC   transcriptase and the molecules are completed to forms that can be
CC   amplified and then expressed in a suitable host-vector system.
CC   The sequence covers all 176 difft base substitutions, most of which
CC   occurred singularly in any given mutant.
CC   See also P80575.
SQ   Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match   1.8%; Score 39; DB 1; Length 204;
Best Local Similarity 13.4%; Pred. No. 6.03e-07;
Matches   16; Conservative   56; Mismatches   47; Indels   0; Gaps   0;

Db   72 gaayycdchvgccgymrttthyrrmrbnvyrdynrsdaaawyccyrrsvkydccynach 131
               || ::|:|:|| || :: ::::| |:: :: ||:.:: :||:::
Qy   46 GAGTCTGCTCGGCGGTGGTGGCCGAGGAAGGAGGAGAGAACGATCGCGGAGCAGGGCGCCC 105

Db   132 hddhyvvybbbvynvhnhnncnccbnnhvchnvhbnnhrnwayvrhdarrddvhccvvch 190
               :::::::|:: :|: |::::|:|:|::|::
Qy   106 GAACTCCGGCCGGGCCGCCATGCCGCCGGCCAGCCGAGACTACGGCAAGTACCTCGCA 164

RESULT 5
ID   T76405 standard; DNA; 178 BP.
AC   T76405;
DT   15-SEP-1997 (first entry)
DE   Human endothelin-1 antisense oligonucleotide.
KW   Asthma; airway epithelium; adenosine free; cystic fibrosis;
KW   chronic obstructive pulmonary disease; bronchitis; ss.

```
OS   Synthetic.
PN   WO9640162-A1.
PD   19-DEC-1996.
PF   06-JUN-1996; U09306.
PR   07-JUN-1995; US-474497.
PA   (UVEC-) UNIV EAST CAROLINA.
PI   Metzger WJ, Nyce JW;
DR   WPI; 97-051871/05.
PT   Treatment of airway diseases such as asthma - by topically applying
PT   adenosine-free antisense oligo:nucleotide to airway epithelium of
PT   subject
PS   Claim 5; Page 38; 71pp; English.
CC   A method for treating airway disease in a subject has been produced,
CC   which involves the topical administration of an essentially adenosine
CC   free antisense oligonucleotide (ON) to the airway epithelium of the
CC   subject. The present sequence is an antisense oligonucleotide specific
CC   for the human endothelin-1, targeted at the initiation codon. The
CC   method can be used to treat airway diseases such as cystic fibrosis,
CC   asthma, chronic obstructive pulmonary disease, bronchitis and other
CC   airway diseases characterised by an inflammatory response. By
CC   eliminating adenosine from the antisense ON, its liberation upon
CC   antisense degradation is prevented, thereby preventing adenosine-
CC   induced bronchoconstriction in patients with hyper-reactive airways.
SQ   Sequence  178 BP;        0 A;    52 C;    46 G;    32 T;

Query Match             1.7%;  Score 37;  DB 32;  Length 178;
  Best Local Similarity   35.38%;
  Matches   36;  Conservative  36;  Mismatches 29;  Indels  1;  Gaps  1;

Db         70  cbgcgcctbgbctgc-tgtttctggbgctccttggcbbgccbcbbbcbbcbbcgbgbgbbb 128
                |:|||||::|:||:|||||||||::|:|| |::|| ||:::: :||:||:::
Qy        253  CGCCGCCTCCCGCGCCATGTTCCATTGCCCTTCCTGGGGTCTGGGACTGGGCCAGGTGGTCTG 312

Db        129  btctgbgcbbbtbbtccbttcgtgbbbbbbgggbtcbbbb 170
                |:||:|| :::: ||:|| ||:||:||::: : |::|
Qy        313  CAGCATCGCTCGTTCCTGTACTTTCGAGCGCAGATGGATCC 354


RESULT 6
ID   T76405 standard; DNA; 178 BP.
AC   T76405;
DT   15-SEP-1997  (first entry)
DE   Human endothelin-1 antisense oligonucleotide.
KW   Asthma; airway epithelium; adenosine free; cystic fibrosis;
KW   chronic obstructive pulmonary disease; bronchitis; ss.
OS   Synthetic.
PN   WO9640162-A1.
PD   19-DEC-1996.
PF   06-JUN-1996; U09306.
PR   07-JUN-1995; US-474497.
PA   (UVEC-) UNIV EAST CAROLINA.
PI   Metzger WJ, Nyce JW;
DR   WPI; 97-051871/05.
PT   Treatment of airway diseases such as asthma - by topically applying
PT   adenosine-free antisense oligo:nucleotide to airway epithelium of
PT   subject
PS   Claim 5; Page 38; 71pp; English.
CC   A method for treating airway disease in a subject has been produced,
CC   which involves the topical administration of an essentially adenosine
CC   free antisense oligonucleotide (ON) to the airway epithelium of the
CC   subject. The present sequence is an antisense oligonucleotide specific
CC   for the human endothelin-1, targeted at the initiation codon. The
CC   method can be used to treat airway diseases such as cystic fibrosis,
CC   asthma, chronic obstructive pulmonary disease, bronchitis and other
CC   airway diseases characterised by an inflammatory response. By
CC   eliminating adenosine from the antisense ON, its liberation upon
CC   antisense degradation is prevented, thereby preventing adenosine-
CC   induced bronchoconstriction in patients with hyper-reactive airways.
SQ   Sequence  178 BP;        0 A;    52 C;    46 G;    32 T;

Query Match             1.7%;  Score 37;  DB 32;  Length 178;
  Best Local Similarity   34.2%;   Pred. No. 6.19e-06;
  Matches   26;  Conservative  33;  Mismatches 16;  Indels  1;  Gaps  1;
```

```
Db        102  tggcbbgccbcbbbcbgcbgbgbbbtcbtgbgcbbb-tbbtccbbtctgbbbbbbbg 160
                |||::|:|:|:||||::|:::.:::|:||:|||::::||:|||: :| |:::::|
Cp        137  TGGCCGGCGCATGGCGGCGGCGCCGGAGTTCGGCGCGCCCTCCGGCGATCGTTCTCTC 78

Db        161  ggbtcbbbbbcctccc 176
                :||::::::|| |||
Cp         77  CCTTCCTCGGCCACCC 62


RESULT 7
ID   T76170 standard; DNA; 88 BP.
AC   T76170;
DT   12-SEP-1997  (first entry)
DE   Human IL3 receptor antisense oligonucleotide.
KW   Asthma; airway epithelium; adenosine free; cystic fibrosis;
KW   chronic obstructive pulmonary disease; bronchitis; interleukin; ss.
OS   Synthetic.
PN   WO9640162-A1.
PD   19-DEC-1996.
PF   06-JUN-1996; U09306.
PR   07-JUN-1995; US-474497.
PA   (UVEC-) UNIV EAST CAROLINA.
PI   Metzger WJ, Nyce JW;
DR   WPI; 97-051871/05.
PT   Treatment of airway diseases such as asthma - by topically applying
PT   adenosine-free antisense oligo:nucleotide to airway epithelium of
PT   subject
PS   Example 5; Page 28; 71pp; English.
CC   A method for treating airway disease in a subject has been produced,
CC   which involves the topical administration of an essentially adenosine
CC   free antisense oligonucleotide (ON) to the airway epithelium of the
CC   subject. The present sequence is an antisense oligonucleotide specific
CC   for the human IL3 receptor, targeted at the initiation codon. The method
CC   can be used to treat airway diseases such as cystic fibrosis, asthma,
CC   chronic obstructive pulmonary disease, bronchitis and other airway
CC   diseases characterised by an inflammatory response. By eliminating
CC   adenosine from the antisense ON, its liberation upon antisense
CC   degradation is prevented, thereby preventing adenosine-induced
CC   bronchoconstriction in patients with hyper-reactive airways.
SQ   Sequence  88 BP;        0 A;    22 C;    36 G;     5 T;

Query Match             1.6%;  Score 35;  DB 32;  Length 88;
  Best Local Similarity   44.6%;   Pred. No. 6.11e-05;
  Matches   37;  Conservative  22;  Mismatches 24;  Indels  0;  Gaps  0;

Db          1  gcbggbgbcbggcgcbggcgcbggbggcbgcgtgbgccbbbgbgbggbgccbctcgggbbcg 60
                ||:|| |||:|| |:|| ||:||:|||:|||||::||||||||||||||:|||:||||
Cp        243  GCCGGAGCCGGTGCAGAAGGCGGCGGGGTGCAGCGGGGTCGCGTGTGGACGCCGGACCCGGGGCCG 184

Db         61  cbgctccggbbgcbgbcbgbg 83
                |:||:|| ::|:|:||:|:|:|
Cp        183  CTGCCCATCTCCTCCGAGCTGCG 161


RESULT 8
ID   Q70469 standard; DNA; 114 BP.
AC   Q70469;
DT   07-APR-1995  (first entry)
DE   Generic DNA sequence to generate a random TSAR peptide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concatenated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.
OS   Synthetic.
FH   Key            Location/Qualifiers
FT   misc_feature    55..60
FT                   /tag= a
FT                   /note= "this sequence represents 'z'; z can be a
FT                   sequence of 6,9 or 12 nucleotides (see
FT                   comments)"
PN   WO9418818-A.
AC   WO9418818-A.
PD   18-AUG-1994.
PF   01-FEB-1994; U00977.
```

PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
DR WPI; 94-279739/34.
PI Fowlkes DM, Kay BK;
PT Identifying proteins or peptides(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70469 is a generic DNA sequence used to generate random TSAR peptide
CC This generic formula can be represented as follows: X(TGC)(NNB)10-
CC (TGC)(NNB)6Z(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
CC sites (X is not the same as Y) that are not specified further. This
CC sequence generates peptides that are cloverleaf in structure. Other
CC generic sequences are shown in Q70465-68. Other specific peptides
CC generated by these generic sequences are shown in R65150-54. TSARs are
CC concatenated heterofunctional proteins or peptides, comprising at least
CC two functional regions - a binding domain with affinity for a ligand and
CC a second effector peptide portion that is chemically or biologically
CC active.They may further comprise a linker peptide between the 2 domains.
CC The oligonucleotides are also designed so that the expressed peptide
CC contains 2 or 4 cysteine residues positioned in, or flanking, the
CC unpredicted or variant residues. These residues confer some degree of
CC conformational rigidity to the peptides. The TSARs or compsns. comprising
CC a TSAR binding domain can be used in vivo to deliver a chemically or
CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC or enzyme, to the specific target or on the cell. They can also replace
CC the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC and therefore circumvent the need for complex methods of hybridoma
CC formation or in vivo antibody production. The TSARs are easily
CC characterised and have designed activity allowing direct and rapid
CC detection in a screening process.
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Query Match   1.6%; Score 36; DB 12; Length 114;
Best Local Similarity 6.5%; Pred. No. 1.96e-05;
Matches 7; Conservative 31; Mismatches 70; Indels 0; Gaps 0;

Db   6 bnnbnnbnnbnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnnbnnnnnnnbnn 65
                 :: :||| : :   :|| :|:
Qy 185 GGCCCCGGCGTCCCACACGAGGTCCGCTGCACCCCGGCCCTTCTGCACCGGCTCCGGCG 244

Db   66 btgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbtg 113
             : :||:|| :   :   :| :|||
Qy 245 CCGCCACCCGCCGCCGCTCCCGCTCCATGTTCCTGGCCCTCCTGGGGCTG 292

RESULT 9
ID Q70465 standard; DNA; 114 BP.
AC Q70465;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR petide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key   Location/Qualifiers
FT misc_feature 55..60
FT   /*tag=   a
FT   /note= "this sequence represents 'z'; z can be a
FT   sequence of 6, 9 or 12 nucleotides (see
FT   comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65150 and R65151.

PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows: X(NNB)6(TGC)(NNB)12(NNB)14(TGC)(NNB)3Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active.They may further comprise a linker
CC peptide between the 2 domains.The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues.These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match   1.6%; Score 36; DB 12; Length 114;
Best Local Similarity 4.5%; Pred. No. 1.96e-05;
Matches 5; Conservative 33; Mismatches 74; Indels 0; Gaps 0;

Db   3 bnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnnbtgcnnbnnbnnnnnnnnn 62
                        :|||
Qy 197 CCACACGAGGTCCGCTGCACCCCGGCCCTTCTGCACCGGCTCCGGCGCCACCCGCC 256

Db   63 bnnbnnbnnbnnbnnbtgcnnbnnbnnbnnbtgcnnbnnbnnb 114
              : :||
Qy 257 GCCTCCGCGTCCATGTTCCTGGCCCTCCTGGGGCTGGTGGACTGGGCCAGGTGG 308

RESULT 10
ID Q70467 standard; DNA; 114 BP.
AC Q70467;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR petide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key   Location/Qualifiers
FT misc_feature 55..60
FT   /*tag=   a
FT   /note= "this sequence represents 'z'; z can be a
FT   sequence of 6, 9 or 12 nucleotides (see
FT   comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65153.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(TGC)(NNB)1Y. X

```
CC   and Y are flanking restriction sites (X is not the same as Y) that are
CC   not specified further. Other generic sequences are shown in Q70466-68.
CC   Other specific peptides generated by these generic sequences are shown in
CC   R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC   comprising at least two functional regions - a binding domain with
CC   affinity for a ligand and a second effector peptide portion that is
CC   chemically or biologically active.They may further comprise a linker
CC   peptide between the 2 domains.The oligonucleotides are also designed so
CC   that the expressed peptide contains 2 or 4 cysteine residues positioned
CC   in, or flanking, the unpredicted or variant residues. These residues
CC   confer some degree of conformational rigidity to the peptides. The TSARs
CC   or compsns. comprising a TSAR binding domain can be used in vivo to
CC   deliver a chemically or biologically active moiety, eg. metal ion,
CC   radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC   cell. They can also replace the function of macromolecules, eg.
CC   monoclonal or polyclonal antibodies and therefore circumvent the need for
CC   complex methods of hybridoma formation or in vivo antibody production.
CC   The TSARs are easily characterised and have designed activity allowing
CC   direct and rapid detection in a screening process.
SQ   Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match
Best Local Similarity   1.6%;     Score 34;  DB 12;  Length 114;
                        3.7%;     Pred. No. 1.89e-04;
Matches    4;  Conservative 32;   Mismatches 72;   Indels 0;   Gaps 0;

Db    3  bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbtgcnnbnnnnnnnn        62
Qy  200  CACGAGGGTCCGCTGCACCCGCGCCTTCGCACCGGCGTCCGGCGCCACCCGCGCC  259

Db   63  bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbtg       110
Qy  260  TCCCGCCTCCATGTTCCTGGCCCCTCCTGGGCGTGGGACTGGGCCAGGTG      307

RESULT  11
ID   Q70468 standard; DNA; 114 BP.
AC   Q70468;
DT   05-APR-1995  (first entry)
DE   Generic DNA sequence to generate a random TSAR petide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concatenated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    55..60
FT                   /tag= a
FT                   /note= "this sequence represents 'z'; z can be a
                            sequence of 6, 9 or 12 nucleotides (see
                            comments)"
     WO9418318-A.
PD   18-AUG-1994.
PF   01-FEB-1994; U00977.
PR   01-FEB-1993; US-013416.
PR   30-DEC-1993; US-176500.
PR   31-JAN-1994; US-189331.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM, Kay BK;
DR   WPI; 94-279739/34.
DR   P-PSDB; R65154.
PT   Identifying proteins or peptide(s) which bind a ligand - by
PT   screening a recombinant vector library expressing fusion proteins
PT   comprising a binding domain and an effector domain
PS   Disclosure; Page 35; 255pp; English.
CC   Q70468 is a generic DNA sequence used to generate random TSAR (Totally
CC   Synthetic Affinity Reagents) peptides.This generic formula can also be
CC   represented as follows: X(NNB)11(TGC)(NNB)6z(NNB)10Y. X
CC   and Y are flanking restriction sites (X is not the same as Y) that are
CC   not specified further. Other generic sequences are shown in Q70466-68.
CC   Other specific peptides generated by these generic sequences are shown in
CC   R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC   comprising at least two functional regions - a binding domain with
CC   affinity for a ligand and a second effector peptide portion that is
CC   chemically or biologically active.They may further comprise a linker
```

```
CC   peptide between the 2 domains.The oligonucleotides are also designed so
CC   that the expressed peptide contains 2 or 4 cysteine residues positioned
CC   in, or flanking, the unpredicted or variant residues. These residues
CC   confer some degree of conformational rigidity to the peptides. The TSARs
CC   or compsns. comprising a TSAR binding domain can be used in vivo to
CC   deliver a chemically or biologically active moiety, eg. metal ion,
CC   radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC   cell. They can also replace the function of macromolecules, eg.
CC   monoclonal or polyclonal antibodies and therefore circumvent the need
CC   for complex methods of hybridoma formation or in vivo antibody
CC   production. The TSARs are easily characterised and have designed activity
CC   allowing direct and rapid detection in a screening process.
SQ   Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match
Best Local Similarity   1.6%;     Score 34;  DB 12;  Length 114;
                        3.6%;     Pred. No. 1.89e-04;
Matches    4;  Conservative 33;   Mismatches 75;   Indels 0;   Gaps 0;

Db    3  bnnbnnbnnbnnbnnbnnbnnbnnbnnbtgcnnbnnbnnbnnbnnnnnnnn       62
Qy  191  GGCGTCCACACCAGAGGTCCGGTCGCACCCGCGCCTTCGCACCGGCGTCCGGCGCCGCCA  250

Db   63  bnnbnnbnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnnb    114
Qy  251  CCCGCCGGCCTCCCGCTCCATGTTCCTGGCCCCTCCTGGGCGTGGGACTGGGCC     302

RESULT  12
ID   Q70467 standard; DNA; 114 BP.
AC   Q70467;
DT   05-APR-1995  (first entry)
DE   Generic DNA sequence to generate a random TSAR petide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concatenated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    55..60
FT                   /tag= a
FT                   /note= "this sequence represents 'z'; z can be a
                            sequence of 6, 9 or 12 nucleotides (see
                            comments)"
PN   WO9418318-A.
PD   18-AUG-1994.
PF   01-FEB-1994; U00977.
PR   01-FEB-1993; US-013416.
PR   30-DEC-1993; US-176500.
PR   31-JAN-1994; US-189331.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM, Kay BK;
DR   WPI; 94-279739/34.
DR   P-PSDB; R65153.
PT   Identifying proteins or peptide(s) which bind a ligand - by
PT   screening a recombinant vector library expressing fusion proteins
PT   comprising a binding domain and an effector domain
PS   Disclosure; Page 35; 255pp; English.
CC   Q70467 is a generic DNA sequence used to generate a random TSAR (Totally
CC   Synthetic Affinity Reagents) peptides.This generic formula can also be
CC   represented as follows: X(NNB)16(TGC)(NNB)1z(NNB)16(TGC)(NNB)1Y. X
CC   and Y are flanking restriction sites (X is not the same as Y) that are
CC   not specified further. Other generic sequences are shown in Q70466-68.
CC   Other specific peptides generated by these generic sequences are shown in
CC   R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC   comprising at least two functional regions - a binding domain with
CC   affinity for a ligand and a second effector peptide portion that is
CC   chemically or biologically active.They may further comprise a linker
CC   peptide between the 2 domains.The oligonucleotides are also designed so
CC   that the expressed peptide contains 2 or 4 cysteine residues positioned
CC   in, or flanking, the unpredicted or variant residues. These residues
CC   confer some degree of conformational rigidity to the peptides. The TSARs
CC   or compsns. comprising a TSAR binding domain can be used in vivo to
CC   deliver a chemically or biologically active moiety, eg. metal ion,
CC   radioisotope, peptide, toxin or enzyme, to the specific target or on the
```

CC   cell. They can also replace the function of macromolecules, eg.
CC   monoclonal or polyclonal antibodies and therefore circumvent the need for
CC   complex methods of hybridoma formation or in vivo antibody production.
CC   The TSARs are easily characterised and have designed activity allowing
CC   direct and rapid detection in a screening process.
SQ   Sequence 114 BP;   0 A;   2 C;   2 G;   2 T;

Query Match   1.6%;   Score 35; DB 12; Length 114;
Best Local Similarity 3.7%; Pred. No. 6.11e-05;
Matches   4; Conservative 33; Mismatches 72; Indels 0; Gaps 0;

Db   6 bnnbnnbnnbnnbnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnn 65
Cp   205 CTCGTGTGGACGCCGGGCCGCTGCCATCTCCTCCGAGCTGCGCAGGTACTTGCCGTA 146

Db   66 bnnbnnbnnbnnbnnbnnbnnbnnbnnbtgcnnbnnbnnbtgcnnb 114
Cp   145 GTCTCGGCTGCCCGCGCATGGCGCGGGGCCGCCGGGAGTTCGGGCGCCCCT 97

RESULT 13
ID   Q70469 standard; DNA; 114 BP.
AC   Q70469;
DT   07-APR-1995   (first entry)
DE   Generic DNA sequence to generate a random TSAR peptide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concateneated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.
OS   Synthetic.
FH   Key   Location/Qualifiers
FT   misc_feature   55..60
FT   /tag= a
FT   /note= "this sequence represents 'z'; z can be a
FT   sequence of 6, 9 or 12 nucleotides (see
FT   comments)"
PN   WO9418318-A.
PD   18-AUG-1994.
PF   01-FEB-1994; U00977.
PR   01-FEB-1993; US-013416.
PR   30-DEC-1993; US-176500.
PR   31-JAN-1994; US-189331.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM, Kay BK;
DR   WPI; 94-279739/34.
PT   Identifying proteins or peptide(s) which bind a ligand - by
PT   screening a recombinant vector library expressing fusion proteins
PT   comprising a binding domain and an effector domain
PS   Disclosure; Page 35; 255pp; English.
CC   Q70469 is a generic DNA sequence used to generate a random TSAR peptide.
CC   This generic formula can be represented as follows: X(TGC)(NNB)10-
CC   (TGC)(NNB)52(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are not specified further.
CC   sites (X is not the same as Y) that are not specified further. Other
CC   sequence generates peptides that are cloverleaf in structure. Other
CC   generic sequences are shown in Q70465-68. Other specific peptides
CC   generated by these generic sequences are shown in R65150-54. TSARs are
CC   concatenated heterofunctional proteins or peptides, comprising at least
CC   two functional regions - a binding domain with affinity for a ligand and
CC   a second effector peptide portion that is chemically or biologically
CC   active.They may further comprise a linker peptide between the 2 domains.
CC   The oligonucleotides are also designed so that the expressed peptide
CC   contains 2 or 4 cysteine residues positioned in, or flanking, the peptide
CC   unpredicted or variant residues. These residues confer some degree of
CC   conformational rigidity to the peptides. The TSARs or compsns. comprising
CC   a TSAR binding domain can be used in vivo to deliver a chemically or
CC   biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC   or enzyme, to the specific target or on the cell. They can also replace
CC   the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC   and therefore circumvent the need for complex methods of hybridoma
CC   formation or in vivo antibody production. The TSARs are easily
CC   characterised and have designed activity allowing direct and rapid
CC   detection in a screening process.
SQ   Sequence 114 BP;   0 A;   4 C;   4 G;   4 T;

Query Match   1.6%;   Score 35; DB 12; Length 114;
Best Local Similarity 7.2%; Pred. No. 6.11e-05;
Matches   8; Conservative 30; Mismatches 73; Indels 0; Gaps 0;

Db   2 gcnnbnnbnnbnnbnnbnnbnnbtgcnnbnnbnnbnnbnnnnnnn 61
Cp   184 GCTGCCCATCTCCTCCGAGCTGCGCAGGTACTTGCCGTAGTCTCGGCTGGCCGCCGGCGCAT 125

Db   62 nbnnbtgcnnbnnbnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbt 112
Cp   124 GGGCGGCGGCCGGGAGTTCGGGCCCTGCTCCGCGATCGTTCTCCCTT 74

RESULT 14
ID   Q70465 standard; DNA; 114 BP.
AC   Q70465;
DT   05-APR-1995   (first entry)
DE   Generic DNA sequence to generate a random TSAR peptide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concateneated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.
OS   Synthetic.
FH   Key   Location/Qualifiers
FT   misc_feature   55..60
FT   /tag= a
FT   /note= "this sequence represents 'z'; z can be a
FT   sequence of 6, 9 or 12 nucleotides (see
FT   comments)"
PN   WO9418318-A.
PD   18-AUG-1994.
PF   01-FEB-1994; U00977.
PR   01-FEB-1993; US-013416.
PR   30-DEC-1993; US-176500.
PR   31-JAN-1994; US-189331.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM, Kay BK;
DR   WPI; 94-279739/34.
DR   P-PSDB; R65150 and R65151.
PT   Identifying proteins or peptide(s) which bind a ligand - by
PT   screening a recombinant vector library expressing fusion proteins
PT   comprising a binding domain and an effector domain
PS   Disclosure; Page 35; 255pp; English.
CC   Q70465 is a generic DNA sequence used to generate a random TSAR (Totally
CC   Synthetic Affinity Reagents) peptides.This generic formula can also be
CC   represented as follows: X(NNB)6(TGC)(NNB)11Z(NNB)14(TGC)(NNB)3Y. X
CC   and Y are flanking restriction sites (X is not the same as Y) that are
CC   not specified further. Other generic sequences are shown in Q70466-68.
CC   Other specific peptides generated by these generic sequences are shown in
CC   R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC   comprising at least two functional regions - a binding domain with
CC   affinity for a ligand and a second effector peptide portion that is
CC   chemically or biologically active.They may further comprise a linker
CC   peptide between the 2 domains.The oligonucleotides are also designed so
CC   that the expressed peptide contains 2 or 4 cysteine residues positioned
CC   in, or flanking, the unpredicted or variant residues. These residues
CC   confer some degree of conformational rigidity to the peptides. The TSARs
CC   or compsns. comprising a TSAR binding domain can be used in vivo to
CC   deliver a chemically or biologically active moiety, eg. metal ion,
CC   radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC   cell. They can also replace the function of macromolecules, eg.
CC   monoclonal or polyclonal antibodies and therefore circumvent the need
CC   for complex methods of hybridoma formation or in vivo antibody
CC   production. The TSARs are easily characterised and have designed
CC   activity allowing direct and rapid detection in a screening process.
SQ   Sequence 114 BP;   0 A;   2 C;   2 G;   2 T;

Query Match   1.6%;   Score 34; DB 12; Length 114;
Best Local Similarity 3.6%; Pred. No. 1.89e-04;
Matches   4; Conservative 33; Mismatches 75; Indels 0; Gaps 0;

Db   3 bnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnnbnnnnnnn 62
Cp   217 GTGCAGCGGACCCTCGTGTGGACGCCGGGCCGCTGCCATCTCCTCCGAGCTGCGCAG 158

Job time : 463 secs.

```
Db        63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbtgcnnbnnbnnb 114
             : .. : .. : .. : .. : .. : .. : .. : || : .. : ..
Cp       157 GTACTTGCCGTAGTCTCGGCTGGCCCGGCCCGGCCATGGCCGCGGCCCCGGAGTTC 106

RESULT   15
ID   Q70468 standard; DNA; 114 BP.
AC   Q70468;
DT   05-APR-1995  (first entry)
DE   Generic DNA sequence to generate a random TSAR petide library.
KW   TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW   effector domain; concateneated heterofunctional protein; linker;
KW   direct; rapid; detection; screening; treatment; generic; ss.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_feature    55..60
FT                   /tag= a
FT                   /note= "this sequence represents 'Z'; z can be a
FT                   sequence of 6, 9 or 12 nucleotides (see
FT                   comments)"
RN   WO9418318-A.
PD   18-AUG-1994.
PF   01-FEB-1994; U00977.
PR   01-FEB-1993; US-013416.
PR   30-DEC-1993; US-176500.
PR   31-JAN-1994; US-189331.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM, Kay BK;
DR   WPI; 94-279739/34.
DR   P-PSDB; R65154.
PT   Identifying proteins or peptide(s) which bind a ligand - by
PT   screening a recombinant vector library expressing fusion proteins
PT   comprising a binding domain and an effector domain
PS   Disclosure; Page 35; 255pp; English.
CC   Q70468 is a generic DNA sequence used to generate random TSAR (Totally
CC   Synthetic Affinity Reagents) peptides.This generic formula can also be
CC   represented as follows: X(NNB)11(TGC)(NNB)6Z(NNB)7(TGC)(NNB)10Y. X
CC   and Y are flanking restriction sites (X is not the same as Y) that are
CC   not specified further. Other generic sequences are shown in Q70466-68.
CC   Other specific peptides generated by these generic sequences are shown in
CC   R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC   comprising at least two functional regions - a binding domain with
CC   affinity for a ligand and a second effector peptide portion that is
CC   chemically or biologically active.They may further comprise a linker
CC   peptide between the 2 domains.The oligonucleotides are also designed so
CC   that the expressed peptide contains 2 or 4 cysteine residues positioned
CC   in, or flanking, the unpredicted or variant residues. These residues
CC   confer some degree of conformational rigidity to the peptides. The TSARs
CC   or compsns. comprising a TSAR binding domain can be used in vivo to
CC   deliver a chemically or biologically active moiety, eg. metal ion,
CC   radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC   cell. They can also replace the function of macromolecules, eg.
CC   monoclonal or polyclonal antibodies and therefore circumvent the need
CC   for complex methods of hybridoma formation or in vivo antibody
CC   production. The TSARs are easily characterised and have designed activity
CC   allowing direct and rapid detection in a screening process.
SQ   Sequence 114 BP;   0 A;   2 C;   2 G;   2 T;

Query Match             1.6%;  Score 34;  DB 12;  Length 114;
Best Local Similarity   3.6%;  Pred. No. 1.89e-04;
  Matches    4;  Conservative   33;  Mismatches  75;  Indels   0;  Gaps   0;

Db         3 bnnbnnbnnbnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnnnnnn 62
             : .. : .. : .. : .. : || : .. : .. : .. : .. : ..
Cp       187 GCCGGCTGCCCATCTCCTCCGAGCTGCGCAGGTACTTGCCGTAGTCTCGGCTGGCCCGGGCCG 128

Db        63 bnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnnbnnb 114
             : .. : .. : || : .. : .. : .. : .. : ..
Cp       127 CATGGCGCGGCGCCCGGAGTCGGGCCGCCCCTGCTCCGGCGATCGTTCTCTCCC 76
```

Search completed: Sat Nov 14 03:20:42 1998

```
****************************************************
 __    __  ____   ___   ___   ___  __   __
|  \  /  ||  _ \ / __| |  _ \ / __| |  | |  |
|   \/   || |_) |\___ \| |_) || |   |  |_|  |
| |\  /| ||  __/  ___) ||  _ < | |   |   _   |
|_| \/ |_||_|    |____/ |_| \_\\___| |__| |__|  (TM)
****************************************************
```

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

****

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sat Nov 14 00:39:32 1998;   MasPar time 2910.85 Seconds
                                            1538.313 Million cell updates/sec

Molecular output not generated.

Title: >US-08-989-362-1
Description: (1-2191) from US08989362.seq
Perfect Score: 2191
N.A. Sequence: 2191   1 GCCAGGACCTCTGTGAACCG..........TTTTGGTACTTAAAAATGGC 2191
Comp: CGGTCCTGGAGACACTTGGC..........AAAACCATGAATTTTTACCG

Scoring table: TABLE default
               Gap 6

Nmatch STD :   Dbase 0: Query 0

Searched: 552174 seqs, 1021863385 bases x 2

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database: embl55
               1:em_ba 2:em_htg 3:em_hum1 4:em_hum2 5:em_in 6:em_om
               7:em_or 8:em_ov 9:em_pat 10:em_pl 11:em_ro
               genbank107
               12:gb_ba 13:gb_htg 14:gb_in 15:gb_om 16:gb_ov 17:gb_pat
               18:gb_ph 19:gb_pl 20:gb_pr1 21:gb_pr2 22:gb_ro 23:gb_st
               24:gb_sts 25:gb_sy 26:gb_un 27:gb_vi

Statistics: Mean 12.129; Variance 6.937; scale 1.749
               Mean 12.129; Variance 6.937; scale 1.749

               Pred. No. is the number of results predicted by chance to have a
               score greater than or equal to the score of the result being printed,
               and is derived by analysis of the total score distribution.

                         SUMMARIES
```
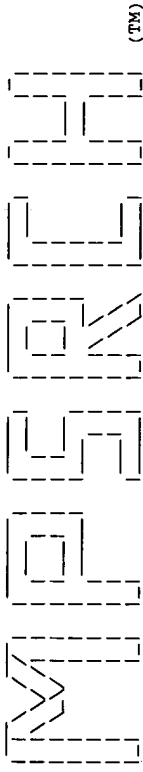
Result No.	Score	Query % Match	Length	DB	ID	Description	Pred. No.
1	2152	98.2	2299	22	AF053713	Mus musculus osteoprot	0.00e+00
2	2150	98.1	2225	22	AF019048	Mus musculus receptor	0.00e+00
3	2070	94.5	2237	22	AF013170	Mus musculus TNF-relat	0.00e+00
4	951	43.4	951	22	AB008426	Mus musculus mRNA for	0.00e+00
5	848	38.7	2271	21	AF053712	Homo sapiens osteoprot	0.00e+00
6	842	38.4	2201	21	AF019047	Homo sapiens receptor	0.00e+00
7	715	32.6	1823	21	AF013171	Homo sapiens TNF-relat	0.00e+00
8	55	2.5	10772	14	AF012089	Drosophila melanogaste	6.20e-15
9	52	2.4	10772	14	AF012089	Drosophila melanogaste	3.84e-13
c 10	47	2.1	7218	17	I66494	Sequence 14 from paten	3.22e-10
c 11	41	1.9	7218	17	I66494	Sequence 14 from paten	7.68e-07
c 12	39	1.8	215	17	I28278	Sequence 5 from patent	9.41e-06
13	40	1.8	74371	21	AC005369	Homo sapiens chromosom	2.70e-06
c 14	37	1.7	74371	21	AC005369	Homo sapiens chromosom	1.10e-04
c 15	35	1.6	215	17	I28278	Sequence 5 from patent	1.21e-03

						Description	
c 16	34	1.6	216021	21	HUAC004787	Homo sapiens Chromosom	3.91e-03
c 17	32	1.5	216021	21	HUAC004787	Homo sapiens Chromosom	3.91e-02
18	30	1.4	4050	14	DDPROKING	D.discoideum gene for	3.63e-01
19	30	1.4	103338	19	AC002294	Arabidopsis thaliana c	3.63e-01
20	30	1.4	161708	13	AC005295	Homo sapiens chromosom	3.63e-01
21	28	1.3	201	17	A10158	S.griseus gene for pre	3.10e+00
22	28	1.3	201	17	A10161	Synthetic DNA for prep	3.10e+00
23	28	1.3	201	17	A10159	S.griseus gene for pre	3.10e+00
c 24	28	1.3	201	17	A10162	Synthetic DNA for prep	3.10e+00
c 25	28	1.3	402	21	SOU60998	Saguinus oedipus clone	3.10e+00
26	28	1.3	421	24	G08367	human STS CHLC.GATA6B	3.10e+00
c 27	28	1.3	565	17	E04076	gDNA encoding envelope	3.10e+00
28	29	1.3	1265	19	DDITFIID	D.discoideum TFIID mRN	1.07e+00
29	29	1.3	1448	19	IBWUN185	I.batatas DNA for woun	1.07e+00
c 30	29	1.3	1596	14	PFAGGMPREP	Plasmodium chabaudi DN	1.07e+00
c 31	29	1.3	2368	17	A12164	S.tuberosum wun1-85 se	1.07e+00
c 32	29	1.3	2368	17	A01611	S.tuberosum recombinan	1.07e+00
33	29	1.3	2368	17	A20297	wun1 and flancking reg	1.07e+00
34	29	1.3	2696	19	DDIPRKNGPK	Dictyostelium discoide	1.07e+00
c 35	28	1.3	84440	19	AB010075	Arabidopsis thaliana g	3.10e+00
36	28	1.3	90271	13	AC004899	Homo sapiens clone DJ0	3.10e+00
37	29	1.3	104071	13	AC004138	*** SEQUENCING IN PROG	1.07e+00
38	28	1.3	111390	19	AC000348	Genomic sequence for A	1.07e+00
39	29	1.3	113803	21	AC004460	Homo sapiens PAC clone	1.07e+00
40	28	1.3	118086	19	ATF6H11	Arabidopsis thaliana D	3.10e+00
41	29	1.3	125520	13	AC005139	*** SEQUENCING IN PROG	1.07e+00
c 42	28	1.3	149316	13	AC005008	Homo sapiens clone GS0	3.10e+00
c 43	28	1.3	168827	13	HS61688	Human DNA sequence ***	3.10e+00
44	28	1.3	182705	13	AC005084	Homo sapiens clone RG2	3.10e+00
c 45	28	1.3	198552	19	ATAP22	Arabidopsis thaliana D	3.10e+00

```
                         ALIGNMENTS

RESULT      1
LOCUS       AF053713     2299 bp    mRNA            ROD       09-MAY-1998
DEFINITION  Mus musculus osteoprotegerin ligand mRNA, complete cds.
ACCESSION   AF053713
NID         g3057147
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2299)
  AUTHORS   Lacey,D.L., Timms,E., Tan,H.-L., Kelley,M.J., Dunstan,C.R.,
            Burgess,T., Elliott,R., Colombero,A., Elliott,G., Scully,S.,
            Hsu,H., Sullivan,J., Hawkins,N., Davy,E., Capparelli,C., Eli,A.,
            Qian,Y.-X., Kaufman,S., Sarosi,I., Shalhoub,V., Senaldi,G., Guo,J.,
            Delaney,J. and Boyle,W.J.
  TITLE     Osteoprotegerin ligand is a cytokine that regulates osteoclast
            differentiation and activation
  JOURNAL   Cell 93 (2), 165-176 (1998)
  MEDLINE   98227661
REFERENCE   2  (bases 1 to 2299)
  AUTHORS   Boyle,W.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAR-1998) Department of Cell Biology, Amgen, Inc.,
            One Amgen Center Drive, Thousand Oaks, California 91320, USA
FEATURES             Location/Qualifiers
     source          1..2299
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             170..1120
                     /function="regulates osteoclast differentiation and
                     activation"
                     /codon_start=1
                     /product="osteoprotegerin ligand"
                     /db_xref="PID:g3057148"
                     /translation="MRRASRDYGKYLRSSEEMSGPGVPHEGPLHPAPSAPAPAPPPA
                     ASRSMFLALLGLGQVVCSIALFLYFRAQMDDPNRISEDSTHCFYRILRLHENAGLQD
                     STLESEDTLPDSCRRMQAFQGAVQKELQHIVGPQRFSGAPAMMEGSWLDVAQRGKPE
                     AQPFAHLTINRASIPSGSHKVTLSSWYHDRGWAKISNMTLSNGKLRVNQDGFYLYAN
```

```
                    ICFRHHETSGSVPDTYLQLMVYYVVKTSIKIPSSHNLMKGGSTKNWSGNSEPHFYSINV
                    GGFFKLRAGEEISIQVSNPSLLDPDQDATYFGAFKVQDID

BASE COUNT    641 a    494 c    541 g    623 t
ORIGIN

Query Match           98.2%;   Score 2152;  DB 22;  Length 2299;
Best Local Similarity  99.5%;  Pred. No. 0.00e+00;
Matches  2177;  Conservative  0;  Mismatches  7;  Indels  3;  Gaps  3;

Db   46  GCCAGGACCTCGTGAACCGTCGGGGCGGGGGCCCCTGCCCGGGAGTCTTGCTCGGCGG   105
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1  GCCAGGACCTCGTGAACCGTCGGGGCGGGGGCCCCTGCCCGGGAGTCTTGCTCGGCGG    60

Db  106  TGGGTGGCCGGAGGAGGAGAACGATCGGCGGAGCAGGGCCCCGAACTCCGGGGCCCG   165
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   61  TGGGTGGCCGGAGGAGGAGAACGATCGGCGGAGCAGGGCCCCGAACTCCGGGGCCCG   120

Db  166  CGCCATGCGCGGGCCCAGACTACGGCAAGTACCTGCGCAGCTCGGAGGAGGATGGG   225
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  121  CGCCATGCGCGGGCCCAGACTACGGCAAGTACCTGCGCAGCTCGGAGGAGGATGGG   180

Db  226  CAGCGGCCCCGGCCGTCCCAACGAGGTCCGCTGCACCCCGCCTTCTGCACCGGCTCC   285
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  181  CAGCGGCCCCGGCCGTCCCAACGAGGTCCGCTGCACCCCGCCTTCTGCACCGGCTCC   240

Db  286  GGGCCGCCCACCCGCCGCCTCCCGCCTCCATGTTCCTGGCCCTCTGGGGCTGGGACTGGG   345
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  241  GGGCCGCCCACCCGCCGCCTCCCGCCTCCATGTTCCTGGCCCTCTGGGGCTGGGACTGGG   300

Db  346  CCAGGTGGTCTGCAGCATCGCTCTGTCTCCTGTACTTCGAGGACACTCCTGACTCCTAACAG   405
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  301  CCAGGTGGTCTGCAGCATCGCTCTGTCTCCTGTACTTCGAGGACACTCCTGACTCCTAACAG   360

Db  406  AATATCAGAAGACAGCACTCACTGCTTTTATAGAATCCTGAGAATCCTGAGAAAACGCAGG   465
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  361  AATATCAGAAGACAGCACTCACTGCTTTTATAGAATCCTGAGAATCCTGAGAAAACGCAGG   420

Db  466  TTTGCAGGACTCGACTCTGGAGAGTGAAGACACACTACCTGACTCCTGCAGGAGGATGAA   525
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  421  TTTGCAGGACTCGACTCTGGAGAGTGAAGACACACTACCTGACTCCTGCAGGAGGATGAA   480

Db  526  ACAAGCCTTTCAGGGGCCGTGCAGAAGGAACTCGAACACATTGTGGGGCCACAGCCGCTT   585
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  481  ACAAGCCTTTCAGGGGCCGTGCAGAAGGAACTCGAACACATTGTGGGGCCACAGCCGCTT   540

Db  586  CTCAGGAGCTCCAGCTATGATGGAAGGCTCATGGTTGGATGTGGCCCAGCGAGGCAAGCC   645
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  541  CTCAGGAGCTCCAGCTATGATGGAAGGCTCATGGTTGGATGTGGCCCAGCGAGGCAAGCC   600

Db  646  TGAGGCCCAGCCCATTGCACACTAAGGGTTAACCTCACCATCAATGCTGCCAGCATCCCA   705
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  601  TGAGGCCCAGCCCATTGCACACTAAGGGTTAACCTCACCATCAATGCTGCCAGCATCCCA   660

Db  706  TAAAGTCACTCGTCCTCTTGGTACCACGACGCTGGGCCAAGATCTCTTCAGCTGATGGT   765
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  661  TAAAGTCACTCGTCCTCTTGGTACCACGACGCTGGGCCAAGATCTCTTCAGCTGATGGT   720

Db  766  GTTAAGCAACGGAAAACTAAGGGTTAACCAAGATGCCTTCATTAACCTGTACGCCAACAT   825
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  721  GTTAAGCAACGGAAAACTAAGGGTTAACCAAGATGCCTTCATTAACCTGTACGCCAACAT   780

Db  826  TTGCCTTTCGGCATCATGAAACATCGGGAAGCCTACAGACTATCTTCAGCTGATGGT   885
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  781  TTGCCTTTCGGCATCATGAAACATCGGGAAGCCTACAGACTATCTTCAGCTGATGGT   840

Db  886  GTATGTCGTTAAAAACCAGCATCAAAATCCCAAGTTCTCATAACCTGATGAAAGGAGGAG   945
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  841  GTATGTCGTTAAAAACCAGCATCAAAATCCCAAGTTCTCATAACCTGATGAAAGGAGGAG   900

Db  946  CACGAAAAACTGGTCGGGCAATTCTGAATTCCACTTTATTCCATAAATGTTGGGGGATT  1005
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  901  CACGAAAAACTGGTCGGGCAATTCTGAATTCCACTTTATTCCATAAATGTTGGGGGATT   960

Db 1006  TTTCAAGCTTCCGAGCTGGTGAAGAAATAGCATTCAGGTGTCCAACCCTTCCCTGCTGGA  1065
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  961  TTTCAAGCTTCCGAGCTGGTGAAGAAATAGCATTCAGGTGTCCAACCCTTCCCTGCTGGA  1020

Db 1066  TCCGGATCAAGATGCGACGTGGTGGTGAAGAAATAGCATTCAAAGTCAAGGACTGAGACTC  1125
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1021  TCCGGATCAAGATGCGACGTGGTGGTGAAGAAATAGCATTCAAAGTCAAGGACTGAGACTC  1080

Db 1126  ATTTCGTGGAACATTAGCATCGATGTCCTAGATGTTGTTGGAAACTTCTTAAAAAATGGATG  1185
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1081  ATTTCGTGGAACATTAGCATCGATGTCCTAGATGTTGTTGGAAACTTCTTAAAAAATGGATG  1140

Db 1186  ATGTCTATACATGTGTAAGACTACTAAGAGACACATGGCCCACGGTGTATGAGAACTCACAGC  1245
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1141  ATGTCTATACATGTGTAAGACTACTAAGAGACACATGGCCCACGGTGTATGAGAACTCACAGC  1200

Db 1246  CCTCTCTTGAGCCTGTACAGGTTGTGTATATGTAAAGTCCATAGGTGATGTTTAGATTC  1305
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1201  CCTCTCTTGAGCCTGTACAGGTTGTGTATATGTAAAGTCCATAGGTGATGTTTAGATTC  1260

Db 1306  ATGGTGATTACACAACGGTTTACAATTTGTAATGATTTCCTA - GAATTGAACCAGATT  1364
         ||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Qy 1261  ATGGTGATTACACAACGGTTTACAATTTGTAATGATTTCCTAAGAATTGAACCAGATT  1320

Db 1365  GGGAGAGTATTCCGATGCTTATGGAAAACTTACACGTGCTATGGAAGGGGTCACAG  1424
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1321  GGGAGAGTATTCCGATGCTTATGGAAAACTTACACGTGCTATGGAAGGGGTCACAG  1380

Db 1425  TCTCTGGGTCTAACCCCTGGTGCCACATGTGCCACTGAGAACCTTGAAATTAAGAGGATGCCAT  1484
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1381  TCTCTGGGTCTAACCCCTGGTGCCACATGTGCCACTGAGAACCTTGAAATTAAGAGAGGTGCCAT  1440

Db 1485  GTCATTGCAAAGAAATGATAGTGTGAAGGGTTAAGTTCTTTGAATTGTTACATTGCGCT  1544
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1441  GTCATTGCAAAGAAATGATAGTGTGAAGGGTTAAGTTCTTTGAATTGTTACATTGCGCT  1500

Db 1545  GGGACCTGCAAATAAGTTCTTTTTTCTAATGAGAGAAAAATATATGTATTTTTATA  1604
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1501  GGGACCTGCAAATAAGTTCTTTTTTCTAATGAGAGAAAAATATATGTATTTTTATA  1560

Db 1605  TAATGTCTAAAGTTATATTTCAGGTGTAATGTTTCTGTGTCAAAGTTTTGTAAATTATAT  1664
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1561  TAATGTCTAAAGTTATATTTCAGGTGTAATGTTTCTGTGTCAAAGTTTTGTAAATTATAT  1620

Db 1665  TTGTGCTATAGTATTTGATTCAAAATATTTAAAAATGTCTCACTGTTGACATATTTAATG  1724
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1621  TTGTGCTATAGTATTTGATTTGATTCAAAATATTTAAAAATGTCTCACTGTTGACATATTTAATG  1680

Db 1725  TTTTAAATGTACAGATGTATTTAACTGGTGCACTTTCCCCTGAAGGTACTCGTA  1784
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1681  TTTTAAATGTACAGATGTATTTAACTGGTGCACTTTGTAATTTCCCCTGAAGGTACTCGTA  1740

Db 1785  GCTAAGGGGCAGAATACTGTTCTGTGACCACATGTAGTGTTATTTCTTTATTCTTTTT  1844
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1741  GCTAAGGGGCAGAATACTGTTCTGTGACCACATGTAGTGTTATTTCTTTATTCTTTTT  1800

Db 1845  AACTTAAGAGTCTTCAGACTTGTCAAGACTATCTGTCAAGCAAAATAAATAAAAAATA  1904
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1801  AACTTAAGAGTCTTCAGACTTGTCAAGACTATCTGTCAAGCAAAATAAATAAAAAATA  1860

Db 1905  AAATGAATACCTTGAATAATAAGTAGGATGTTGGCACCAGGTGCCTTTCAAATTTAGAA  1964
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1861  AAATGAATATCTTGAATAATAAGTAGGATGTTGGTGCCACCAGGTGCCTTTCAAATTTAGAA  1920

Db 1965  GCTAATTGACTTTAGGAGCTGACATAGCCAAAAAGGATACATAATGGCTACTGAAA-TC  2023
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Qy 1921  GCTAATTGACTTTAGGAGCTGACATAGCCAAAAAGGATACATAATAGGCTACTGAAAATC  1980

Db 2024  TGTCAGGAGTATTATGCAATTATTGAACAGTGGCTCTTTTTACACAGGTGTCTTTTTTTTACAAATTG  2083
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1981  TGTCAGGAGTATTATGCAATTATTGAACAGGTGTCTTTTTTTACACAGAGCTACAAATTG  2040

Db 2084  TAAATTTGTTCTTTTTTTTTCCCATAGAAAAATGTACTATAGTGTTATCAGCCAAAAACA  2143
```

```
Qy   2041  TAAATTTTGTTCTTTTTTTTTCCCATAGAAAAATGTACTATAGTTTTATCAGCAAAAAAACA  2100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2144  ATCCACTTTTTAATTTAGTGAAAGTATTTTTATTATATACTGTACAATAAAAAGCATTGTCTC  2203
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Qy   2101  ATCCACTTTTTAATTTAGTGAAAGTATTTTTATTATATACTGTACAATAAAAAGCATTGTTTC  2160

Db   2204  TGAATGTTAATTTTTGGTACAAAAAA  2230
           ||||| |  |  ||||||||||  ||||
Qy   2161  TGAATGGCA-TTTTTGGTACTTAAAA  2186
```

Query Match   98.1%; Score 2150; DB 22;   Length 2225;
Best Local Similarity 99.5%; Pred. No. 0.00e+00;
Matches 2176; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

```
Db     13  GCCAGGACCTCGTGAACCGGTCGGGGCGGGGGCCGCCTGCCCGGGAGTCGCTCGGCGG  72
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1  GCCAGGACCTCGTGAACCGGTCGGGGCGGGGGCCGCCTGCCCGGGAGTCGCTCGGCGG  60

Db     73  TGGGTGGCCGAGGAAGGGAGAGAACGATCGCGGAGCAGGGCCCGGCGCCCGAACTCCGGGCGCCGCC  132
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61  TGGGTGGCCGAGGAAGGGAGAGAACGATCGCGGAGCAGGGCCCGGCGCCCGAACTCCGGGCGCCG  120

Db    133  CGCCATGGCGCCGGCGGGCCAGCCGAGACTACGGCAAGTACCTGCGCAGCTCGGAGGAGATGGG  192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121  CGCCATGGCGCCGGCGGGCCAGCCGAGACTACGGCAAGTACCTGCGCAGCTCGGAGGAGATGGG  180

Db    193  CAGCGGCCCCGGCGGGTCCGCTGCCCGCTCCATGTTCCTGGCCCTCCTGGGGTGGGACTGGG  252
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181  CAGCGGCCCCGGCGGGTCCGCTGCCCGCTCCATGTTCCTGGCCCTCCTGGGGTGGGACTGGG  240

Db    253  GGGCCCGGCCACCCGCCGCCCGCTCCATGTTCCTGGCCCTCCTGGGGTGGGACTGGG  312
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241  GGGCCGCCGCCACCCGCCGCCCGCTCCATGTTCCTGGCCCTCCTGGGGTGGGACTGGG  300

Db    313  CCAGGTGTCTGCAGCATCGCTCGTTCTGTCCTGTACTTTCGAGCGCAGATGATCCTAACAG  372
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301  CCAGGTGGTCTGCAGCATCGCTCGTTCTGTCCTGTACTTTCGAGCGCAGATGATCCTAACAG  360

Db    373  AATATCAGAAGACAGCACTCACTGCTTTTATAGAATCCTGAGACTCCATGAAAACCAGA  432
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    361  AATATCAGAAGACAGCACTCACTGCTTTTATAGAATCCTGAGACTCCATGAAAACGCAGG  420

Db    433  TTTGCAGGACTCGACTCTGGAGAGTGAAGACACACTACCTGACTCCTGCAGGAGGATGAA  492
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421  TTTGCAGGACTCGACTCTGGAGAGTGAAGACACACTACCTGACTCCTGCAGGAGGATGAA  480

Db    493  ACAAGCCTTTCAGGGGCCGTGCAGAAGGAACTCCAACACATTGTGGGGCCACAGCGCTT  552
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    481  ACAAGCCTTTCAGGGGCCGTGCAGAAGGAACTGCAACACATTGTGGGGCCACAGCGCTT  540

Db    553  CTCAGGAGCTCCAGCTATGATGGAAGGCTCATGGTTGGATGTGGCCCAGCGAGGCAAGCC  612
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    541  CTCAGGAGCTCCAGCTATGATGGAAGGCTCATGGTTGGATGTGGCCCAGCGAGGCAAGCC  600

Db    613  TGAGGCCAGCCATTTGCACACCTCACCATCAATGCTGCCAGCATCCATCGGGTTCCCA  672
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    601  TGAGGCCCAGCCATTTGCACACCTCACCATCAATGCTGCCAGCATCCATCGGGTTCCCA  660

Db    673  TAAAGTCACTCTGTCCTCTTGGTACCAGATCGAGGCTGGCCAAGATCTCTAACATGAC  732
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    661  TAAAGTCACTCTGTCCTCTTGGTACCAGATCGAGGCTGGCCAAGATCTCTAACATGAC  720

Db    733  GTTAAGCAACGGAGAAAACTAAGGGTAACCAAGATGGCTTCTATTACCTGTACGCCAACAT  792
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    721  GTTAAGCAACGGAGAAAACTAAGGGTAACCAAGATGGCTTCTATTACCTGTACGCCAACAT  780

Db    793  TTGCTTTCGGCATCATGAAACATCGGGAAGCGTACCTACAGACTATCTTCAGCTGATGGT  852
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    781  TTGCTTTCGGCATCATGAAACATCGGGAAGCGTACCTACAGACTATCTTCAGCTGATGGT  840

Db    853  GTATGTCGTTAAAACCAGCATCAAAATCCAAGTTCTATAAACCTGATGAAAGAGAGGGAG  912
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    841  GTATGTCGTTAAAACCAGCATCAAAATCCAAGTTCTCATAAACCTGATGAAAGAGAGGGAG  900

Db    913  CACGAAAAACTGGTCGGGCAATTCGAATTCCACTTTTATTCCATAAAGTGTGGGGATT  972
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    901  CACGAAAAACTGGTCGGGCAATTCGAATTCCACTTTTATTCCATAAAGTGTGGGGATT  960

Db    973  TTTCAAGCTCCGAGCTGGTGAAGAAATTAGCATTCAGGTGTCCAACCCTTCCCTGCTGGA  1032
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    961  TTTCAAGCTCCGAGCTGGTGAAGAAATTAGCATTCAGGTGTCCAACCCTTCCCTGCTGGA  1020

Db   1033  TCCGGATCAAGATGCGACGTACTTTGGGGCTTCAAAGTTCAGGACATAGACTGAGACTC  1092
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1021  TCCGGATCAAGATGCGACGTACTTTGGGGCTTCAAAGTTCAGGACATAGACTGAGACTC  1080

Db   1093  ATTTCGTTGGAACATTAGCATGGATGTCCTAGATGTTTGGAAACTTCTTAAAAATGGATG  1152
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1081  ATTTCGTTGGAACATTAGCATGGATGTCCTAGATGTTTGGAAACTTCTTAAAAATGGATG  1140

Db   1153  ATGTCTATACATGTGTAAGACTACTAAGAGACATGGCCCACGGTGTATGGAAACTCACAGC  1212
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1141  ATGTCTATACATGTGTAAGACTACTAAGAGACATGGCCCACGGTGTATGGAAACTCACAGC  1200

Db   1213  CCTCTCTTGAGCCTGTACAGGTTGTGTATATGTAAGTCCATAGGTGATGTTAGATTC  1272
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1201  CCTCTCTTGAGCCTGTACAGGTTGTGTATATGTAAGTCCATAGGTGATGTTAGATTC  1260
```

```
ORGANISM    Mus musculus
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2237)
AUTHORS     Wong,B.R., Rho,J., Arron,J., Robinson,E., Orlinick,J., Chao,M.,
            Kalachikov,S., Cayani,E., Bartlett,F.S. III, Frankel,W.N., Lee,S.Y.
            and Choi,Y.
TITLE       TRANCE is a novel ligand of the tumor necrosis factor receptor
            family that activates c-Jun N-terminal kinase in T cells
JOURNAL     J. Biol. Chem. 272 (40), 25190-25194 (1997)
MEDLINE     97460112
REFERENCE   2  (bases 1 to 2237)
AUTHORS     Wong,B.R., Rho,J., Arron,J., Lee,S.Y., Robinson,E. and Choi,Y.
TITLE       Direct Submission
JOURNAL     Submitted (09-JUL-1997) Howard Hughes Medical Institue, The
            Rockefeller University, 1230 York Ave., New York, NY 10021, USA
FEATURES             Location/Qualifiers
     source          1..2237
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="14"
                     /map="near Rb1"
                     /cell_line="T cell hybridoma KMls-8.3.5"
     CDS             142..1092
                     /note="TNF-related ligand"
                     /codon_start=1
                     /product="TRANCE"
                     /db_xref="PID:g2411498"
                     /translation="MRRASRDYGKYLRSSEEMSSGPGVPHEGPLHPAPSAPAPAPPPA
                     ASRSMFLALLGLGLQQVVCSIALPLYFRAQMDPNRISEDSTHCFYRILRLHENAGLQD
                     STLESEDTLPDSCRRMKQAFQGAVQKELQHIVGPQRFSGAPAMMEGSWLDVAQRGKPE
                     AQPPAHLTINAASIPSGSHKVTLSSWYHDRGWAKISNMTLSNGKLRVNQDGFYLYAN
                     ICFRHHETSGSVPTDYLQLMVYYVVKTSIKIPSSHNLMKGGSTKNWSGNSFHFYSINV
                     GGFFKLRAGEEISIQVSNPSLLDPDQDATYFGAFKVQDID"
BASE COUNT      636 a    470 c    519 g    612 t
ORIGIN
```

```
Query Match            Best Local Similarity  94.5%;   Score 2070;  DB 22;  Length 2237;
                                              98.88%;   Pred. No. 0.00e+00;
Matches 2162; Conservative 0; Mismatches 14; Indels 13; Gaps 10;
```

```
Db   22 GCCAGGACCTTGTGAACCGGTCGGGGGCGTCGGGGGCCG--TGGC--GGAGTCTGCTCGGCGG   77
         ||||||||| ||||||||||||||||||| |||||||||     |    |||||||||||||||
Qy    1 GCCAGGACCTCGTGAACCGGTCGGGGCGGGGGGGCCGTCGGCGGGAGTCTGCTCGGCGG   60

Db   78 TGGGTGGCCCGAGAAGGAGAACGATCGCGGAGCAGGGGCCCGAACTCCGGGCGGCCG  137
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   61 TGGGTGGCCCGAGAAGGAGAACGATCGCGGAGCAGGGGCCCGAACTCCGGGCGGCCG  120

Db  138 CGCCATGCGCCGGGCCAGCCGACTACGGCAAGTACCTGCGCAGCTCGGAAGAGATGGG  197
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  121 CGCCATGCGCCGGGCCAGCCGACTACGGCAAGTACCTGCGCAGCTCGGAAGAGATGGG  180

Db  198 CAGCGGCCCCGGCGTCCCACACGGAGGTCCGCTGCACCCCGCGCCCTTCTGCACCGCTCC  257
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  181 CAGCGGCCCCGGCGTCCCACACGGAGGTCCGCTGCACCCCGCGCCCTTCTGCACCGCTCC  240

Db  258 GGCGCCGCCACCCCGCGCTCCCCGCTCCTTCTGGGGCTGGAGCTGGGGACTGGG  317
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  241 GGCGCCGCCACCCCGCGCTCCCCGCTCCTTCTGGGGCTGGAGCTGGGGACTGGG  300

Db  318 CCAGGTGGTCTGCAGCATCGCTCTGTTCCTGTACTTCAGGGCGCAGATGGATCCTAACAG  377
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  301 CCAGGTGGTCTGCAGCATCGCTCTGTTCCTGTACTTCAGGGCGCAGATGGATCCTAACAG  360

Db  378 AATATCAGAAGACAGCACTCACTGCTTTTATAGAATCCTGAGACTCCATGAAAACGCAGG  437
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  361 AATATCAGAAGACAGCACTCACTGCTTTTATAGAATCCTGAGACTCCATGAAAACGCAGG  420

Db  438 TTTGCAGGACTCGCACTCTGGAGAGTGAAGACACTACCTGACTCCTGCAGGAGGATGAA  497
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  421 TTTGCAGGACTCGCACTCTGGAGAGTGAAGACACTACCTGACTCCTGCAGGAGGATGAA  480
```

```
Db 1273 ATGGTGATTACCAACAACGGTTTTACAATTTGTAATGATTCCTA-GAATTGAACCAGATT 1331
         ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Qy 1261 ATGGTGATTACCAACAACGGTTTTACAATTTGTAATGATTCCTAAGAATTGAACCAGATT 1320

Db 1332 GGGAGAGGTATTCCGATGCTTATGAAAAACTTACACGTGAGCTATGGAAGGGGTCACAG 1391
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1321 GGGAGAGGTATTCCGATGCTTATGAAAAACTTACACGTGAGCTATGGAAGGGGTCACAG 1380

Db 1392 TCTCTGGTCTAACCCCTGGACATGTGCCACTGAGAACCTGAAATTAAGAGGATGCCAT 1451
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1381 TCTCTGGTCTAACCCCTGGACATGTGCCACTGAGAACCTGAAATTAAGAGAGATGCCAT 1440

Db 1452 GTCATTGCAAAGAAATGATAGTGGAAGGGTTAAGTTCTTTGAATTGTTACATTGCGCT 1511
         |||||||||||||||||||||| ||||||||||||||||||||||||| ||||||||||
Qy 1441 GTCATTGCAAAGAAATGATAGTGTGAAGGGTTAAGTTCTTTGAATTGTTCACATTGCGCT 1500

Db 1512 GGGACCTGCAAATAAGTTCTTTTTTTCTAATGAGGAGAGAAAAATATATGTATTTTTATA 1571
         ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Qy 1501 GGGACCTGCAAATAAGTTCTTTTTTTCTAATGAGGAGAGAAAAATATGTATTTTTATA 1560

Db 1572 TAATGTCTAAAGTTATATTCCAGGTGTAATGTTTCTGTGCAAAGTTTGTAAATTATAT 1631
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1561 TAATGTCTAAAGTTATATTCCAGGTGTAATGTTTCTGTGCAAAGTTTGTAAATTATAT 1620

Db 1632 TTGTGCTATAGTATTTGATTCAAAATATTAAAAATGTCTCACTGTTGACATATTTAATG 1691
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1621 TTGTGCTATAGTATTTGATTCAAAATATTAAAAATGTCTCACTGTTGACATATTTAATG 1680

Db 1692 TTTTAAATGTACAGATGTATTTAACTGGTGCACTTTGTAATTCCCCTGAAGGTACTCGTA 1751
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1681 TTTTAAATGTACAGATGTATTTAACTGGTGCACTTTGTAATTCCCCTGAAGGTACTCGTA 1740

Db 1752 GCTAAGGGGGCAGAATACTGTTTCTGGTGACCACATGTAGTTTATTCTTATTCTTTTT 1811
         ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Qy 1741 GCTAAGGGGGCAGAATACTGTTTCTGGTGACCACATGTAGTTTATTCTTTTATTCTTTTT 1800

Db 1812 AACTTAATAGAGTCTTCAGACTTGTCAAAACTATGCAAGCAAATAAAATAAATAAAAATA 1871
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1801 AACTTAATAGAGTCTTCAGACTTGTCAAAACTATGCAAGCAAATAAAATAAATAAAAATA 1860

Db 1872 AAATGAATAACCTTGAATAATAAGTAGGATGTTGGTCACCAGTTGCCTTCAAATTTAGAA 1931
         |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Qy 1861 AAATGAATAACTTGAATAATAAGTAGGATGTTGGTCACCAGTTGCCTTCAAATTTAGAA 1920

Db 1932 GCTAATTGACTTTAGGAGCTGACATAGCCAAAAAGGATACATAATAGGCTACTGAAA-TC 1990
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Qy 1921 GCTAATTGACTTTAGGAGCTGACATAGCCAAAAAGGATACATAATAGGCTACTGAAAATC 1980

Db 1991 TGTCAGGAGTATTATTGCAATTATTGAACACAGTGTCTTTTTTTACAAGAGTACAAATTG 2050
         |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Qy 1981 TGTCAGGAGTATTATTGCAATTATTGAACACAGTGTCTTTTTTTACAAGAGTACACAATTG 2040

Db 2051 TAAATTTTGTTTTCTTTTTTGTTTCTTTTTTTCCCATAGAAAATGTACTATAGTTTATCAGCCAAAAAACA 2110
         ||||||||||||||||||| | ||||||||||||||||||||||||||||||||||||||
Qy 2041 TAAATTTTGTTTTCTTTTTTTTCCCATAGAAAATGTACTATAGTTTATCAGCCAAAAAACA 2100

Db 2111 ATCCACTTTTAATTTAGTGAAAGTTATTTTATTATACTGTACAATAAAAGCATTGCTTC 2170
         |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Qy 2101 ATCCACTTTTAATTTAGTGAAAGTTATTTTATTATACTGTACAATAAAAGCATTGTTTC 2160

Db 2171 TGAATGTTAATTTTTGGTACAAAAA 2197
         |||||| | |||||||||||  ||||
Qy 2161 TGAATGGCA-TTTTTGGTACTTAAAA 2186
```

```
RESULT    3
LOCUS       AF013170        2237 bp    mRNA            ROD       20-SEP-1997
DEFINITION  Mus musculus TNF-related ligand TRANCE mRNA, complete cds.
ACCESSION   AF013170
NID         g2411497
KEYWORDS    .
SOURCE      house mouse.
```

```
Db  498  ACAAGCCTTTCAGGGGGCCGTGCAGAAGGAACTGCAACACATTGTGGGCCACAGCGCTT   557
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  481  ACAAGCCTTTCAGGGGGCCGTGCAGAAGGAACTGCAACACATTGTGGGCCACAGCGCTT   540

Db  558  CTCAGGAGCTCCAGCTATGATGGAAGGCTCATGGTTGGATGTGGCCCAGCGAGGCAAGCC  617
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  541  CTCAGGAGCTCCAGCTATGATGGAAGGCTCATGGTTGGATGTGGCCCAGCGAGGCAAGCC  600

Db  618  TGAGGCCCAGCCATTGCACACCTCACCACGATCAATGCTGCCAGCCATCCCATCGGGTTCCCA  677
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  601  TGAGGCCCAGCCATTGCACACCTCACCACGATCAATGCTGCCAGCCATCCCATCGGGTTCCCA  660

Db  678  TAAAGTCACTCTGTCCTCTTGGTACCAAGATGGCTTCTATTACCTGTACGCCAACATGAC  737
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  661  TAAAGTCACTCTGTCCTCTTGGTACCAAGATGGCTTCTATTACCTGTACGCCAACATGAC  720

Db  738  GTTAAGCAACGGAAAACTAAGGGTTAACCAAGATGGCTTCTATTACCTGTACGCCAACAT  797
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  721  GTTAAGCAACGGAAAACTAAGGGTTAACCAAGATGGCTTCTATTACCTGTACGCCAACAT  780

Db  798  TTGCTTTCGGCATCATGAAACATCGGGAAGCGTACCTACAGACTATCTTCAGCTGATGGT  857
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  781  TTGCTTTCGGCATCATGAAACATCGGGAAGCGTACCTACAGACTATCTTCAGCTGATGGT  840

Db  858  GTATGTCGTTAAAAACCAGCATCAAAATCCCAAGTTCTCATAACCTGATGAAAGGAGGGAG  917
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  841  GTATGTCGTTAAAAACCAGCATCAAAATCCCAAGTTCTCATAACCTGATGAAAGGAGGGAG  900

Db  918  CACGAAAAACTGGTCGGGCAATTCTGAATTCCACTTTTATTCCATAAAATGTGGGGGATT  977
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  901  CACGAAAAACTGGTCGGGCAATTCTGAATTCCACTTTTATTCCATAAAATGTGGGGGATT  960

Db  978  TTTCAAGCTCCGAGCTGGTGAAGAAATTAGCATTCAGGTGTCCAACCCTTCCCTGCTGGA  1037
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  961  TTTCAAGCTCCGAGCTGGTGAAGAAATTAGCATTCAGGTGTCCAACCCTTCCCTGCTGGA  1020

Db 1038  TCCGGATCAAGATGCGACGTACTTTGGGCTTCAAAGTTCAGGACATAGACTGAGACTC  1097
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1021  TCCGGATCAAGATGCGACGTACTTTGGGCTTCAAAGTTCAGGACACATAGACTGAGACTC  1080

Db 1098  ATTTCGTGGAACATTAGCATGCATGGATGTCCTAGATGTTTGGAAAACTTCCTAAAAAATGATG  1157
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1081  ATTTCGTGGAACATTAGCATGCATGGATGTTTCCTAGATGTTTGGAAAACTTCCTAAAAAATGATG  1140

Db 1158  ATGTCTATACACATGTGTAAGACTACTAAGAGACTACTAAGAGACATGGCCCACGGTGTATGCTGTATGAGCAGC  1217
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1141  ATGTCTATACACATGTGTAAGACTACTAAGAGACTACTAAGAGACATGGCCCACGGTGTATGCTGTATGAAACTCACAGC  1200

Db 1218  CCTCTCTCTTGAGCCCTGTACAGGTTGTGTATATGTAAAGTCCATAGGTGATGTTTAGATT  1277
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1201  CCTCTCTCTTGAGCC-TGTACAGGTTGTGTATATGTAAAGTCCATAGGTGATGTTTAGATT  1259

Db 1278  CATGGTGATTACACAACGGTTTACAATTTTGTAATGATTCCTA-GAATTCGAACCAGAT  1336
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1260  CATGGTGATTACACAACGGTTTACAATTTTGTAATGATTCCTAAGAATTAAGAACCAGAT  1319

Db 1337  TGGGAGAGGTATTCCGATGCTTATGAAAAACTTACACGTGACTATGGAAGGGGTCACA  1396
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1320  TGGGAGAGGTATTCCGATGCTTATGAAAAACTTACACGTGACTATGGAAGGGGTCACA  1379

Db 1397  GTCTCTGG-TCTAACCCCTGGCACATGTGCCACTGAGAACCTTGAAATTAAGAGGATGCCA  1455
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1380  GTCTCTGGGTCTAACCCCTGGCACATGTGCCACTGAGAACCTTGAAATTAAGAGAGATGCCA  1439

Db 1456  TGTCATTGCATTAGAAATGATAGTGTGAAGGGTAAGGGTTCTTTTGAATTGTTACATTGCGC  1515
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1440  TGTCATTGCATTAGAAATGATAGTGTGAAGGGTAAGGGTTCTTTTGAATTGTTACATTGCGC  1499

Db 1516  TGGGACCTGCAAATAGTTCTTTTTCTAATGAGGAGA-AAAATATATGTATTTTTAT  1573
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1500  TGGGACCTGCAAATAGTTCTTTTTCTAATGAGGAGAGAAAAATATATGTATTTTTTAT  1559

Db 1574  ATAATGCTAAAGTATATTCAGGTGTAATGTTGTTTTCTGTGCAAAGTTTGTAAATTATA  1633
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy 1560  ATAAGTCTAAAGTATATATTCAGGTGTAAAGTTTTCGTCGTGTTGACATATTAAT  1619
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1634  TTTGTGCTATAGTATTTGATTCAAAATATTTAAAAATGTCTCACTGTTGACATATTTAAT  1693
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1620  TTTGTGCTATAGTATTTGATTCAAAATATTTAAAAATGTCTCACTGTTGACATATTTAAT  1679

Db 1694  GTTTTAAATGTACAGATGTATTTAACTGGTGCACTTGTAATTCCCCTGAAGTGACTACTCGT  1753
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1680  GTTTTAAATGTACAGATGTATTTAACTGGTGCACTTGTAATTCCCCTGAAGTGACTACTCGT  1739

Db 1754  AGCTAAGGGGCAGAATACTGTTTCTCGTGACCACATGAGTTATTTCTTTATTCTTT  1813
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1740  AGCTAAGGGGCAGAATACTGTTTCTCGTGACCACATGAGTTATTTCTTTATTCTTT  1799

Db 1814  TAACTTAATAGAGTCTTCAGACTTGTCAAAACTATGCAAGCAAAATAAATAAATAAAAAAT  1873
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1800  TAACTTAATAGAGTCTTCAGACTTGTCAAAACTATGCAAGCAAAATAAATAAATAAAAAAT  1859

Db 1874  AAAAATGAATACCTTGAATAATAAGTAGGATGTTGGTCACCAGGTGCCTTTCAAATTTAGA  1933
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1860  AAAATGAATATCTTGAATAATAAGTAGGATGTTGGTCACCAGGTGCCTTTCAAATTTAGA  1919

Db 1934  AGCTAATTGACTTTAGGAGCTGACATAGCCAAAAGGA-ACATAATAGGCTACTGAAA-T  1991
         |||||||||||||||||||||||||||||||||||| ||||||||||||||||
Qy 1920  AGCTAATTGACTTTAGGAGCTGACATAGCCAAAAGGATACATAATAGGCTACTGAAAAT  1979

Db 1992  CTGTCAGGAGTATTATGCAATTATTGAACAGGTGTCTTTTTTTTACAAGAGCTACAAATT  2051
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1980  CTGTCAGGAGTATTATGCAATTATTGAACAGGTGTCTTTTTTTTACAAGAGCTACAAATT  2039

Db 2052  GTAAATTTGGTTTCTTTTTTTTTCCCATAGAGAAAATGTACTATAGTTTATCAGCCAAAAAA  2111
         |||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Qy 2040  GTAAATTTTG-TTTCTTTTTTTTTTCCCATAGAGAAAATGTACTATAGTTTATCAGCCAAAAAA  2098

Db 2112  CAATCCACTTTTTAATTTAGTGAAAGTTATTTTATTATACTGTACAATAAAAGCATTGTC  2171
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2099  CAATCCACTTTTTAATTTAGTGAAAGTTATTTTATTATACTGTACAATAAAAGCATTGTT  2158

Db 2172  TCTGAATGTTAATTTTTTGGTACAAAAA  2200
         ||||||||    |||||||||
Qy 2159  TCTGAATGGCA-TTTTTGGTACTTAAAA  2186
```

Qy   725 AGCAACGCGAAAACTAAGGGTTAACCAAGATGGCTCTTATTACCTGTACGCGCCAACATTGC 784
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TTTCGGCATCATGAAACATCGGAAGCGTACCTACAGACTATCTTCAGCTGATGGTGTAT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   785 TTTCGGCATCATGAAACATCGGAAGCGTACCTACAGACTATCTTCAGCTGATGGTGTAT 844
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GTCGTTAAAACCAGCATCAAAATCCCAAGTTCTATAACCTGATGAAAGGAGGAGCACG 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   845 GTCGTTAAAACCAGCATCAAAATCCCAAGTTCTATAACCTGATGAAAGGAGGAGCACG 904
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AAAAACTGGTCGGCAATTCGAATTCCACTTTATTCCATAAAATGTGGGGGATTTTTC 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   905 AAAAACTGGTCGGCAATTCGAATTCCACTTTATTCCATAAAATGTGGGGGATTTTTC 964
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AAGCTCCGAGCTGGTGAGAAATTAGCATTCAGGTGTCCAACCCTTCCTGCTGGATCCG 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   965 AAGCTCCGAGCTGGTGAGAAATTAGCATTCAGGTGTCCAACCCTTCCTGCTGGATCCG 1024
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GATCAAGATGCGACGTACTTTGGGGTTTCAAAGTTCAGGACATAGACTGA 951
          ||||||||||||||||||||||||||| |||||||||||||||
Qy 1025 GATCAAGATGCGACGTACTTTGGGGCTTTCAAAGTTCAGGACGACTGA 1075

RESULT 5
LOCUS   AF053712   2271 bp   mRNA   PRI   09-MAY-1998
DEFINITION Homo sapiens osteoprotegerin ligand mRNA, complete cds.
ACCESSION   AF053712
NID   g3057145
KEYWORDS   .
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2271)
AUTHORS   Lacey,D.L., Timms,E., Tan,H.-L., Kelley,M.J., Dunstan,C.R.,
            Burgess,T., Elliott,R., Colombero,A., Elliott,G., Scully,S.,
            Hsu,H., Sullivan,J., Hawkins,N., Davy,E., Capparelli,C., Eli,A.,
            Qian,Y.-X., Kaufman,S., Sarosi,I., Shalhoub,V., Senaldi,G., Guo,J.,
            Delaney,J. and Boyle,W.J.
TITLE   Osteoprotegerin ligand is a cytokine that regulates osteoclast
            differentiation and activation
JOURNAL   Cell 93 (2), 165-176 (1998)
MEDLINE   98227661
REFERENCE   2 (bases 1 to 2271)
AUTHORS   Boyle,W.J.
TITLE   Direct Submission
JOURNAL   Submitted (16-MAR-1998) Department of Cell Biology, Amgen, Inc.,
            One Amgen Center Drive, Thousand Oaks, California 91320, USA
FEATURES   Location/Qualifiers
     source   1..2271
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS   185..1138
                     /codon_start=1
                     /product="osteoprotegerin ligand"
                     /db_xref="PID:g3057146"
                     /function="regulates osteoclast differentiation and
                     activation"
                     /translation="MRRASRDYTKYLRGSEEMGGGPGAPHEGPLHAPPPPAPHQPPAA
                     SRSMFVALLGLGQVYCSVALFFYFRAQMDPNRISEDGTHCIYRILRHENADFQDT
                     TLESQDTKLIPDSCRRIKQAFQGAVQKELQHIVGSQHIRAEKAMVDGSWLDLAKRSKL
                     EAQPFAHLTINATDIPSGSHKVSLSSWYHDRGWAKISNMTFSNGKLIVNQDGFYYLYA
                     NICFRHHETSGDLATEYLQLMVYVTKTSIKIPSSHILMKGGSTKYWSGNSEFHFYSIN
                     VGGFFKLRSGEEISIEVSNPSLLDPDQDATYFGAFKVRDID"
BASE COUNT   658 a   462 c   522 g   629 t
ORIGIN

Query Match   38.7%; Score 848; DB 21; Length 2271;
Best Local Similarity 78.5%; Pred. No. 0.00e+00;
Matches 1642; Conservative 0; Mismatches 380; Indels 69; Gaps 43;

Db   173 GGGGCCGAGCGGCCATGGCGCCGCCCGCCAGCCAGAGACTACCACCAAGTACCTGCGTGGCTCGGAG 232
          |||| |||||||||||||||||||||||||||||||||||||||||||||||

FEATURES   Location/Qualifiers
     source   1..951
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /cell_line="T2"
                     /cell_type="stromal cells"
                     /clone="pOBM291"
                     /clone_lib="pcDL-SRa296"
                     /tissue_type="bone marrow"
     CDS   1..951
                     /codon_start=1
                     /product="osteoclast differentiation factor (ODF)"
                     /db_xref="PID:d1026353"
                     /db_xref="PID:g3041782"
                     /translation="MRRASRDYGKYLRSSEEMGSGPGVPHEGPLHPAPSAPAPAPPPA
                     ASRMFLALLGLGQVVCSIALFLYFRAQMDPNRISEDSTHCFYRLRLHENAGLQD
                     STLESEDTLPDSCRRMKQAFQGAVQKELQHIVGPQRFSGAPAMMEGSWLDVAQRGKPE
                     AQPFAHLTINAASIPSGSHKVTLLSSWYHDRGWAKISNMTLSNGKLRVNQDGFYYLYAN
                     ICFRHHETSGSVPTDYLQLMVYVVKTSIKIPSSHNLMKGGSTKNWSGNSEFHFYSINV
                     GGFFKLRAGEEISIQVSNPSLLDPDQDATYFGAFKVQDID"
     misc_feature   142..213
                     /note="transmenbrane domain"
     misc_feature   454..948
                     /note="Tumor Necrosis Factor(TNF)-like domain"
BASE COUNT   231 a   267 c   248 g   205 t
ORIGIN

Query Match   43.4%; Score 951; DB 22; Length 951;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 951; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db   1 ATGGCGCGGGCCAGCCGAGACTACGGCAAGTACCTGCGCAGCTCGGAGGAGATGGGCAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   125 ATGGCGCGGGCCAGCCGAGACTACGGCAAGTACCTGCGCAGCTCGGAGGAGATGGGCAGC 184
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GGCCCCGGGGTCCCACACGAGGGTCCGCTCCACCCCGCCCCGTCCGCCCCGGCTCCGGCG 120
          |||||| || |||||||||||||||||||||||||||||||||||||||||||||||||
Qy   185 GGCCCCGGCGTCCCACACGAGGGTCCGCTCCACCCCGCCCCGTCCGCCCCGGCTCCGGCG 244
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CGGCCACCCGCGCCCGCTCCCGCCTCCATGTTCCTGGCCCTCCTGGGACTCGGGCCAG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Qy   245 CGGCCACCCGCGCCCGCTCCCGCCTCCATGTTCCTGGCCCTCCTGGGACTGGGCCAG 304
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTGGTCTGCAGCATCGCTCTGTTCCTGTACTTTCGAGCGCAGATGGATCCTAACAGAATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   305 GTGGTCTGCAGCATCGCTCTGTTCCTGTACTTTCGAGCGCAGATGGATCCTAACAGAATA 364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TCAGAAGACAGCACTCACTGCTTTTATAGAATCCTGAGACTCCATGAGAAAACGCAGTTTG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   365 TCAGAAGACAGCACTCACTGCTTTTATAGAATCCTGAGACTCCATGAGAAAACGCAGTTTG 424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGGACTCGGACTCGGAGAGTGAAGACAACACTACCTGACTCCTGCAGGAGGATGAAACAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   425 CAGGACTCGGACTCGGAGAGTGAAGACAACACTACCTGACTCCTGCAGGAGGATGAAACAA 484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCCTTTCAGGGGGCCGTGCAGAAGGAACTGCAACACATTGTGGGGCCCACAGCCGCTTCTCA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   485 GCCTTTCAGGGGGCCGTGCAGAAGGAACTGCAACACATTGTGGGGCCCACAGCCGCTTCTCA 544
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GGAGCTCCAGCTATGATGGAAGGCTCATGGTTGGATGTGGCCCAGCGGGGCAAGCCTGAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   545 GGAGCTCCAGCTATGATGGAAGGCTCATGGTTGGATGTGGCCCAGCGGGGCAAGCCTGAG 604
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GCCCAGCCATTGCACACCTCACCATCAATGCTGCCAGCATCCCATCCGGTTCCCATAAA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   605 GCCCAGCCATTGCACACCTCACCATCAATGCTGCCAGCATCCCATCCGGTTCCCATAAA 664
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GTCACTCTGTTGTCCTCTTGGTACCACGATCGAGGCTGGGCCAAGATCTCTAACATGACGTTA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   665 GTCACTCTGTTGTCCTCTTGGTACCACGATCGAGGCTGGGCCAAGATCTCTAACATGACGTTA 724
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AGCAACGCGAAAACTAAGGGTTAACCAAGATGGCTCTTATTACCTGTACGCCCAACATTGC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Qy  113  GGGCGCCGCCATGCGCCGGGGGCCGAGACTACGGCAAGCTACCTGCCAGCTCGGAG  172
Db  233  GAGATGGGCGGCGGCCCGAGCCCGCACGAGGGCCCCGCGCCCCGCGCCGCCC-GC-  290
Qy  173  GAGACAGAATATCAGAAGATGCCACTCACTGCATTTATAGAGTTTGAGACTTCCATGA  232

Db  291  -CTGCGCCGCCACCAGCCCGCCCCCGCCTCCGCGCTCCGCTCGTGGGCCTCGGGGCTG  349
Qy  233  CCGGCTCCGGCGCGCCGCCACCCGCCGCCCCGCCGCTCCGCGCTCCGCGCTCGGGGGCTG  292

Db  350  GGCTGGGCCAGGTGTCTGCAGGCGTCGCCCGTGTCTTCTATTTCAGAGCGCAGAGTGGAT  409
Qy  293  GGACTATGGAAGGGGTCACAGTCTCGGGTCTAACCCCTGGACGCGCAGATGGAT  352

Db  410  CCTAATGAGATATCAGAAGATGCCACTCACTGCATTTATAGAGTTTGAGACTTCCATGAA  469
Qy  353  CCTAACAGAATATCACTGCTTTATAGAGTTTGAGACTCCATGAA  412

Db  470  AATGCAGATTTCAAGGAGTACAGGACCAAGTACAAAATTAATACCTGATTCA  529
Qy  413  AACGCAGGTTTGCAGGACTCGACTCTGGAGACTGAAGACACAC------TACCTGACTCC  466

Db  530  TGTAGGAGAATTAAACAGGCCTTTCAAGGAGTCGTGCAAAAGGAATTACACACATCGTT  589
Qy  467  TGCAGGAGGATGAAACAGCCTTTCAGGGGCGTCCAGCGTTGGGATGGGGCC  526

Db  590  GGATCACAGCACATCAGAGCCAGAGAAGCCGATGGTGGATGGGCTCATGGTTAGATCTGGCC  649
Qy  527  GGGCCACAGCGCGTTCTCAGGGAGGCTCCAGCTATGGATGATGGGCCTCATGGTTGGGCC  586

Db  650  AAGAGGAGCAAGCTTGAAGCTCAGCCTTTTGCTCACTACTAATGCCACCGACATC  709
Qy  587  CACGGAGGCAAGCCTGAGGCCCAGCCATTTGCACACTCCACCATCAATGCTGCCAGCATC  646

Db  710  CCATCTGGTTCCCATAAAGTGAGTCTGTCCTCTTGGTACCATGGGGGTTGGGCCAAG  769
Qy  647  CCATCGGTTCCCATAAAGTCACTCTGTCCTCTTGGTACCCAGAGGCTCGAGGGCTGGGCCAAG  706

Db  770  ATCTCCAACATGACTTTTAGCAATGGAAAACTAATAGTTAATCAGGATGGCTTTTATTAC  829
Qy  707  ATCTCCAACATGACGTTAAGCAACGGAAAACTAAGGGGTTAACCAAGATGGCTTCTATTAC  766

Db  830  CTGTATGCCAACATTTGCTTTCGACATCATGAAACTTCAGGAGACCTAGCTACAGAGTAT  889
Qy  767  CTGTACGCCAACATTTGCTTTCGGCATCATGAAACCAGCGGAGCGTACCTACAGACTAT  826

Db  890  CTTCAACTAATGGTGTACGTCACTCAACTAAAACCCAAGTCTCATCACCCTG  949
Qy  827  CTTCAGCTGATGGTGTATGTCGTTCAGTTTGGGGCAATTCTGAAAATCCCACTTTATTCCATA  886

Db  950  ATGAAAGGAGGAAGCACCAAGTATTGGTCAGGAATTCCATTTTTATTCCATA  1009
Qy  887  ATGAAAGGAGGAGGAGCACGAAAAGGTCGGGCAAGTTCTGAATTCCACTTTTATTCCATA  946

Db  1010 AACGTTGGTGGATTTTTAAGTTAAGTTACGGTCCTGGAGGAGGAAATCAGCATCGAGGTCTCCCAAC  1069
Qy  947  AAGTTGGGAGATTTTTCAGCTGCGAGGTCTCCAGCGTGCAGGTGTCCCAAC  1006

Db  1070 CCCTCCTTACTCGGATCCGGATCAGGATCGCAACATACTTTGGGCTTTAAAGTTCGAGAT  1129
Qy  1007 CCTTCCCTGATCCCGGATCCGGATCAAGATGCGACGTACTTTCGGGCTTTCAAAGTTCAGGAC  1066

Db  1130 ATAGATTGAGCCCCAGTTTTTGGAGTGTTA---TGTATTTCCTGGATGTTTGGAAACATT  1186
Qy  1067 ATAGACTGAGACTCATTTTCGTGAACATTGTGGAACATTGGCATTTTTCTGAGTGTTGAAAC--T  1124

Db  1187 TTTTAAAACAGGCCAAGAAAGATCTATATAGGTGTATATAGGTGTGAGACTAAGGAGCATGGCCC  1246
Qy  1125 TCTTAAAAAATG---GAT-GATTGCTATACATGTCTAAGACTGTAAGAGAACACGGCCC  1179

Db  1247 CAACGGTACACGACTCAGTATCCATGCTCTTGACCTTGTAGAGAACACGCGTATTTTACAG  1306
Qy  1180 ACGGTGTATGAAACTCACAGCCCCTCTGCTCTTGAGCCTGTGCTACAGGTTGTGTATATGTGTAAAG  1239
```

```
Db  1307 CCAGTGGGAGATGTTAGACTCATGTGTGTTACACAATGTTTTAAATTTGTAATGAA  1366
Qy  1240 TCCATAGTGATGTTAGATTCATGGTGA-TTACACAACGTTTTACAATTTGTAATGAT  1298

Db  1367 TTCCTA-GAATTAAACCAGATTGG--AGCAATTACGG-G--T-TGAC---CTTATGAGA  1415
Qy  1299 TTCCTAAGAATTGAACCAGATTGGGAGAGGTATTCCGATGCTTATGAAAAACTTACACGT  1358

Db  1416 AA-CT--GCATGTGGCTATGGGAG--GGGT-TGGTCCCTGGTCATGTCCCCTTCGCAG  1469
Qy  1359 GAGCTATGGAAGGGGTCACAGTCTCGGGTCTAACCCCTGGACATGTGCCACTGAGAAC  1418

Db  1470 CT-GAAGTGGAGAGGGTGTCATCT-AGGGCAATGAAGGATCATCTGAAGGGGCAAATTC  1527
Qy  1419 CTTGAAATTAAGAAGATGCCATGTCATTGCAAAGAAATGATAGTGTGAAGGTTAAGTTC  1478

Db  1528 TTTTGAATTGTTACATCTGGAACCTGCAAAAA-TACTTTT--CTAATGAGGAGA  1584
Qy  1479 TTTTGAATTGTTACATTGCGCTGGGACCTGCAAATTAAGTTCTTTTTTCTAATGAGGAGA  1538

Db  1585 GAAAA-TATATGTATTTTATATAATAATCTAAAGTTATATTTCAGATGTAATGTTTTCTT  1643
Qy  1539 GAAAATATATGTATTTTATATTTTATATATATGTCTAAAGTTATATTTCAGGTGTAATGTTTTCTG  1598

Db  1644 TGCAAAGTATTGTAAATTATATTGTGTGCTATAGTATTTGATTCAAAATATTTAAAAATGT  1703
Qy  1599 TGCAAAGTTTTGTAAATTTGTGCTATAGTATTTGATTCAAAAATATTTAAAAATGT  1658

Db  1704 CTTGCTCGTTGACATATTTAATGTTTAAATGTCAGACATATTTAACTGTGCACTTTGT  1763
Qy  1659 CTCACTGTTGACATATTTAATGTTTAAATGTTACAGATGTATTTAACTGTGCACTTTGT  1718

Db  1764 AAATTCCCTGGGGAAAACTTGCAGCTAAGGAGGGGGAAAAAATGTTGTT-TCCTAATATC  1822
Qy  1719 AATTCCCCTGAGGTA-CTCGTAGCTAAGGGGCAGAATACTGTTTCTGTGACCACCATG  1777

Db  1823 AAATGCAGTA-TAT-TTCTTCGTT-CTTTTTA-AGTTAATAAGAGAGATTTTTCAGACTTGTCA  1878
Qy  1778 TAGTTTATTTCTTTATTCTTTTAACTTAATAGAGTCTTCAGACTTCAAAACTATGCA  1837

Db  1879 AGCCTG-TGCA-AA-AAAATTAAAATGATGCCTTGAATAATAAGCAGGATGTGGCCA  1934
Qy  1838 AGCAAAATAAATAAATAAATAAAATGAATATCTTGAATAAGTAAGTAGGATGTGTGGTCA  1897

Db  1935 CCAGGTGCCTTTCAAATTTAGAAAGCTAATTGACTTTAGAAAGCTGACCATTGCCAAAAAGG  1994
Qy  1898 CCAGGTGCCTTTCAAATTTAGAAATTTAGAGAGCTAATTGACTTTAGGA-GCTGACATAGCCAAAAAGG  1956

Db  1995 ATACATAATGGGGCCACTGAAA-TCTGTCAAGGAGTAGTATATATAATTGTTGAACAGGTGT-  2052
Qy  1957 ATACATAATAGGCTACTGCAAATCGTCGAGAAAATCGTCAGGAGTATTTATGCAATTATTGAACAGTGTC  2016

Db  2053 -TTTTCCACAAGTGCCGCAAATTGTGTACCTTT--TT-TTTTTTTCAAAATAGAAAAGTTA  2108
Qy  2017 TTTTTTTTACAAGAGCTACAATTGTAAATTTTGTTCTTTTTTCCCATAGAAAATGTA  2076

Db  2109 TTAGTGGTTTATCAGC--AAAA-AGTCCAATTT-AATTTAGTAAATGTTATCTTAT-  2162
Qy  2077 CTA-TAGTTTATCAGCAATCCACTTTTAATTTAGTTAGTGAAGTTATTTTATTA  2135

Db  2163 -ACTGTACAATAAAAACATTGCCTTTGAATGTTAAATTTTTGGTACAAAAA  2212
Qy  2136 TACTGTACAATAAAAGCATTGTTTCTGGATGGCA-TTTTTGGTACTTAAA  2185
```

```
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2201)
AUTHORS     Anderson,D.M., Maraskovsky,E., Billingsley,W.L., Dougall,W.C.,
            Tometsko,M.E., Roux,E.R., Teepe,M.C., DuBose,R.F., Cosman,D. and
            Galibert,L.
TITLE       A homologue of the TNF receptor and its ligand enhance T-cell
            growth and dendritic-cell function
JOURNAL     Nature 390 (6656), 175-179 (1997)
MEDLINE     98032977
REFERENCE   2  (bases 1 to 2201)
AUTHORS     Anderson,D.M., Billingsley,W., Dougall,W., Maraskovsky,E.,
            Cosman,D., DuBose,R. and Galibert,L.
TITLE       Direct Submission
JOURNAL     Submitted (13-AUG-1997) Molecular Biology, Immunex Corp., 51
            University St., Seattle, WA 98101, USA
FEATURES             Location/Qualifiers
     source          1..2201
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /map="13q14"
     gene            1..2201
                     /gene="RANKL"
     CDS             129..1082
                     /gene="RANKL"
                     /note="receptor activator of nuclear factor kappa B
                     ligand"
                     /codon_start=1
                     /product="RANKL"
                     /db_xref="PID:g2612922"
                     /translation="MRRASRDYTKYLRGSEMGGGPGAPHEGPLHAPPPPAPHQPPAA
                     SRSMFVALLGLGQVVCSVALFFYFRAQMDPNRISEDGTHCIYRILRLHENADFQDT
                     TLESQDTKLIPDSCRRIKQAFQGAVQKELQHIVGSQHIRAEKAMVDGSWLDLAKRSKL
                     EAQPFAHLTINATDIPSGSHKVSLSSWYHDRGWAKISNMTFSNGKLIVNQDGFYYLYA
                     NICFRHHETSGDLATEYLQLMVYVKTSIKIPSSHTLMKGGSTKYWSGNSEFHFYSIN
                     VGGFFKLRSGEEISIVSNPSLLDPDQDATYFGAFKVRDID"
BASE COUNT       658 a    429 c    497 g    617 t
ORIGIN
```

```
Query Match          38.4%;  Score 842;  DB 21;  Length 2201;
Best Local Similarity 78.5%;  Pred. No. 0.00e+00;
Matches 1642;  Conservative 0;  Mismatches 380;  Indels 70;  Gaps 44;
```

```
Db   117 GGGCCGAGCCCATGCCGCCGCCGCCGCCAGCAGAGACTACACCAAGTACCTGCGTGGCCTCCGGAG   176
         |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Qy   113 GGGCGCCGCCGCCGCCGGGGCCGGCCCCGGGCCCCGGGCTCGGCCGCTGCCTCCGCAGCTCCGGAG   172

Db   177 GAGATGGGCGGCGGCGCCCGAGCCCCACAGAGGGCCCCCGCACCGCCCGCGCC--GC-         234
         ||||||| |||||| || | |||||||||||||| | ||||| |||| ||||||  ||
Qy   173 GAGATGGGCAGCGGCGCCGGCCCCGGGCGTCCCACACGAGGGTCCGCTGCACCCCGGCCTTCTGCA   232

Db   235 -CTGGCGCCACCAGCCCCGCGCGCCCCCGGCCCTCCATGTTCGTGGCCCTG                 293
          ||| ||| ||| ||||||||||  |||||||||||||||||||||||||||
Qy   233 CCGGCCTCCGGCCGCCACCCGCCCCGGCCGCCTCCGCCTCCGGGGCTCCTG                 292

Db   294 GGGCTGGGCCAGGTTGTCTGCAGCGTCGCCCTGTTCTTCTATTTCAGAGCGCAGATGAT         353
         ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Qy   293 GGGCTGGGCCAGGTGGTCTGCAGCGTGGCCCTGTTCTTCTACTTTCGAGCGCAGATGGAT         352

Db   354 CCTAATAGAAATATCAGAAGATGGCACTCACTGCATTTATAGAATTTTGAGACTCCATGAA       413
         |||| |||||||| |||||||||||| ||||||||||| ||||||||||||||||||||||
Qy   353 CCTAACAGAATATCAGAAGATGGCACTCACTGCATTTATAGAATCCTGAGACTCCATGAA       412

Db   414 AATGCAGATTTTCAAGACACAACTCTGGAGACACAAAGTCAAGATACAAAAATTAATACCTGATTCA   473
         || ||| |||||| ||| |||||| ||||||| ||| |||||||| |||||||| |||| ||
Qy   413 AACGCAGTTTGCAGGACTCGACTCTGGAGAGTGAAGCTGCAAATAAGTTCTTTTTTTCTAATGAGGAG   473

Db   474 TGTAGGAGAATTAAAACAGGCCCTTCAAGGAGCTGTGCAAAAGGAATTACAAACATATCGTT         533
         || |||| ||||| |||||| ||||||||||||| |||||||||||||||||||| |||||||
Qy   467 TGCAGGAGGATGAAACAGCCTTCAGGGGGCCGTCAGGAGAAGGAACTGCAACACATTGTG        526
```

```
Db   534 GGATCACAGCACATCAGACAGCAGAGAAAGCGATGGTGGATGGCTCATGGTGGATGTGGCC   593
Qy   527 GGGCCACAGCGCTTCAGCAGAGCTCCAGCTATGATGGAAGGCTCATGGTGGATGTGGCC    586

Db   594 AAGAGGAGCAAGCTTGAAGCTCAGCCTTCAGCCATTGGCTATCTCACTATTAATGCCACGCACATC   653
Qy   587 CAGCGGAGGCAAGCCTGAGGCCCAGCCATTGCACACCTCACCATCCAATGCTGCCAGCATC   646

Db   654 CCAATCTGGTTCCCATAAAGTGAGTCTGTCTCCTTGGTACCATGATCGGGGTGGGCCAAG   713
Qy   647 CCATCTGGGTTCCCATAAAGTCACTCTGTCCCCTTGGTACCACGATCGAGGCTGGGCCAAG   706

Db   714 ATCTCCAACATGACTTTTAGCAACAATGAAAACTAATAGTTAATCAGATGGCCTTTATTAC   773
Qy   707 ATCTCTAACATGACGTTAAGCAACGGCAAAAACTAAGGGTTAACCAAGATGGCTTCTATTAC   766

Db   774 CTGTATGCCAACATTGCTTTCGGCATCATGAAACTCAGGAGACCTAGCTACAGAGTAT   833
Qy   767 CTGTACGCCAACATTGCTTTCGGCATCATGAAACATCGGGAAGCGTACCTACAGACTAT   826

Db   834 CTTCAACTAATGGTGTACGTCCACTAAACCAGCATCAAAATCCCAAGTTCTCATACCCTG   893
Qy   827 CTTCAGCTGATGGTGTATGTCGTTAAAACCAGCATCAAAATCCCAAGTTCTCATACACCTG   886

Db   894 ATGAAAGGAGGAAGCACCAAGTATTGGTCAGGGAATTCTCATTTTTATTCCATA   953
Qy   887 ATGAAAGGAGGGAGCACGAAAAACTGGTCGGGCAATTCTGAATTCCACTTTTATTCCATA   946

Db   954 AACGGTTGGTGGATTTTTAAGTTACGGTCTCGGAGGAAATCAGCATCGAGGTCTCCAAC   1013
Qy   947 AATGTTGGGGGATTTTTCAGCTCCGAGCTGGTGAAGAAATTAGCATTCAGTGTCCAAC   1006

Db  1014 CCCTCCTTACTGGATCCGGATCAGGATGCAACATACTTGGGGCTTTTAAAGTTCGAGAT   1073
Qy  1007 CCTTCCCTGCCTGGATCCGGATCAAGATGCAACATGGGGCTTTTGGGGCTTTCAAAGTTCAGGAC   1066

Db  1074 ATAGATTGAGCCCCAGTTTTTGGAGTGTTA---TGTATTTCCTGGATGTTTGGAAACATT   1130
Qy  1067 ATAGACTGAGACTCATTTCGTGGAACATTAGCATTAGCATGGTGTCCTAGTATGTTTTGGAAAC--T   1124

Db  1131 TTTTAAAACAGCAAGAAAGATGTATATAGGTGTGTGAGACTACTAAGAGGCATGGCCC   1190
Qy  1125 TCTTAAAAAATG----GAT-GATGTCTATACATGTGTAAGACTACTAAGAGACATGGCCC   1179

Db  1191 CAACGGTACACGACTCAGTATCCATGCTCCTTGACCTTGTAGAGAACACGCGTATTTACCT   1250
Qy  1180 ACGGTGTATGAAACTCCACAGCCTGTCTCTTGAGCCTGTACAGGTTGTGTATATGTAAAG   1239

Db  1251 GCCAGTGGGAGATGTTAGACTCATGGTGTTTACACAATGTTTTTAAATTTTGTAATGA   1310
Qy  1240 TCCA-TAGGTGATGTTAGATTCATGGTGA--TTACACAACGGTTTACAAATTTTGTAATGA   1297

Db  1311 ATTCCTA-GAATTAAACCAGATGG---AGCAATTACGG-G--T-TGAC---CTTATGAG   1359
Qy  1298 TTTCCTAAGAATTGAACCAGATTGGGAGAGGTATTCCGATGCTTATGAAAAACTTACACG   1357

Db  1360 AAA-CT--GCATGTGGCTATGGGAG--GGGT-TGGTCCCTGTCATGTGCCCCTTCGCA   1413
Qy  1358 TGAGCTATGAAGGGGTGCACAGTCTCGGGTCTAACCCCTGGACATGTGCCACTGAGAA   1417

Db  1414 GCT-GAAGTGGAGAGGGTGTCATCT-AGCCCAATTGAAGGATCATCTGAAGGGGCAAATT   1471
Qy  1418 CCTTGAAATTAAGAAGATGCCATGTCATGTCAAAGAAAATGAGTGTAGGGTTAAGTT   1477

Db  1472 CTTTTGAATTGTTACATCATGCTGGAGTCAAGATACAAAAATTAATACCTGATTCA   1528
Qy  1478 CTTTTGAATTGTTACATTGCGCTGGGACCTGCAAATAAGTTCTTTTTTCTAATGAGGAG   1537

Db  1529 AGAAAA-TATATGTATTTTTATATATATCTAAAGTTATATTTCAGATATAATGTTTTCT   1587
Qy  1538 AGAAAAAATATATGTATTTTTATATAATGTCTAAAGTTCTAAAGTTTATATTTCAGTGTAATGTTTTTCT   1597
```

```
CDS             <1..738
                /note="TNF-related ligand"
                /codon_start=1
                /product="TRANCE"
                /db_xref="PID:g2411500"
                /translation="QMDPNRISEDGTHCIYRIRLRLHENADFQDTTLESQDTKLIPDSC
                RRIKQAFQGAVQKELQHIVGSQHIRAEKAMVDGSWLDLAKRSKLEAQPFAHLTINATD
                IPSGSHKVSLSSWYHDRGWGKISNMTFSNGKLIVNQDGFYYLYANICFRHHETSGDLA
                TEYLQLMVYVTKTSIKIPSSHTLMKGGSTKYWSGNSEFHFYSINVGGFFKLRSGEEIS
                IEVSNPSLLDPDQDATYFGAFKVRDID"

BASE COUNT      569 a    305 c    380 g    569 t
ORIGIN

Query Match             32.6%;   Score 715;  DB 21;  Length 1823;
Best Local Similarity   78.2%;   Pred. No. 0.00e+00;
Matches 1450;           Mismatches 339;  Indels 66;  Gaps 41;
                        Conservative 0;
```

```
Db     1  CAGATGGATCCTAATAGAATATCAGAAGATGGCACTCACTGCATTTATAGAATTTGAGA    60
          ||||||||||||||  |||||||||| ||||||||||||  ||||||||||  |||| |||
Qy   344  CAGATGGATCCTAACGAGATATCAGAAGACAGCACTCACTGCTTTTATAGAATCCTGAGA   403

Db    61  CTCCATGAAAATGCAGATTTCAAGACACAACTCTGGAGAGTCAAGATACAAAATTAATA   120
          ||||| ||| |||  || || || |||| |||   |||||||||||||||||||    ||
Qy   404  CTCCATGAAAACGCAGTTTGCAGACTCGACTCGAGGAGAGTGAAGACACAC------TA   457

Db   121  CCTGATTCATGTAGGAGAATTAAACAGGCCTTTCAAGGAGCTGTGCAAAAGGAATTACAA   180
          ||||| || ||| |||||| || || |||| ||||| |||||| |||   ||| |||||||
Qy   458  CCTGACTCCTGCAGGAGGATGAAACAAGCCTTTCAGGGGGCCGTGCAGAAGGAACTGCAA   517

Db   181  CATATCGTTGGATCACAGCACATCAGAGCACACAGAAGCAGAGGATGGCTCATGGTTA    240
          || |||| |||||  ||   || ||||| ||||| | |||  ||  |||||| ||||||
Qy   518  CACATTGTGGGGCCACCAGCCGCTTCCAGGAGCTCCAGCCTTTGCTCATCTCACTATTAATGCC   577

Db   241  GATCTGGCCAAGAGAGCAAGCTTGAAGCTGCAGCCTTTGCTCATCCACTATTAATGCC    300
          |||| |||||| ||||| ||||||||||| ||| |||||||||||||  ||  |||||||
Qy   578  GATGTGGCCCAGCGAGGCAAGCCTGAGCCCAGCCATTGCACCAGTCATAATGCT   637

Db   301  ACCGGACATCCCATCTGGTTCCCATAAAGTGAGTCTGTCCCTTGGTACCGGTACCGGAAGT    360
          ||| || ||||||||||||||||||||||||||  |||| | ||||||||||||||||||
Qy   638  GCCAGCATCCCATCCGGTTCCCATAAAGTCACTCGTGTCCTCTTGGTACCGACGATCGAGGC   697

Db   361  TGGGGTAAGATCCCAACATGACTTTTAGCAATGGAAAACTATAGTAATCAGGATGGC    420
          |||| ||||||||||||||||||||| ||||  ||  |||||||||||||||||||||
Qy   698  TGGGGCCAAGATCCTAACATGACGTTAAGCAACGGAAAAACTAAGGGTTAACCAAGATGGC   757

Db   421  TTTTATTACCTGATGCCAACATTGCTTGCACATCATGAGAATCATGAAACTTCAGGAGACCTAGCT    480
          ||| |||||||||| | ||||||||||||  |||||||  ||||  ||||||||||||||||||
Qy   758  TTCTATTACCTGTACGCCAACATTGCTTGCTTCCGGCATCATGAAACATCGGGAAGCGTACCT   817

Db   481  ACAGAGTATCTTCAACTAATGGTACGTCACTCAAAATCCAGACACAAATCCCAAGTTCT    540
          |||||||||||||||| |||||||||||  ||||||||||||||||||||||||||||||
Qy   818  ACAGACTATCTTCAGCTGATGGTGTATGTCGTTAAAACCAGCATCAAAATCCCAAGTTCT   877

Db   541  CATACCCTGATGAAAGGAGGAGGAGGAGCACCAAGTATTGGTCAGGAATTCCATTTT    600
          ||| ||||||||||| ||||||  ||  || |||||||  || || |||||||||||
Qy   878  CATACCCTGATGAAGGGAGGATCGGGTGGTGAATTCCATTTCTATTCCATCAATGTT   937

Db   601  TATTCCATAAACGTGGTGATTTTTGGGAATTTTCAAGCTCCGGATCAGCATCAGGAG    660
          || ||||||||||||||||  ||| |||  |||||||||| |||| ||  ||||||||
Qy   938  TATTCCATAAATGTTGGGGGATTTTTCAAGCTCCGGATCAGCATCAGGAGAATTAGCATTCAG   997

Db   661  GTCTCCAACCCTCCTTACTGCACATACTTTGGGGGCTTTTAAA    720
          || |||||||||||| || || ||||||||||| ||||||||||
Qy   998  GTGTCCAACCCTTCCCCGCTGCGATCAAGATGCGACGTACTTTGGGGCTTTCAAA   1057

Db   721  GTTCGAGATATAGATTGAGCCCCAGTTTTGCGAGTGTTA---TGTATTCCTCGGATGTTT    777
          |||| ||| | |||    |||||||||||  ||||| ||   |||||||||||||||||||
Qy  1058  GTTCAGGACATAGACTGATCGATCATCTCGTGAACATCGTGAACATTAGCATTAGCATTCAG   1117

Db   778  GGAAACATTTTTAAAACAAGCAAGAAGATGTATAGGTGTGACTACTAAGAG   837
          |||||| || || ||||||||||||||||||||||||||||||||||||||
Qy  1118  GGAAAC--TTCTTAAAAAATG----GAT-GATGTCTATACACTGTAAGACTACTAAGAG   1170
```

```
Db  1588  TTGCAAAGTATTGTAAATATATTGTGCTATAGTATTGATTCAAAATATTAAAAATG   1647
          |||||||  |||||| |||||||||||||||||||||||||||||||||||||||||
Qy  1598  GTGCAAAGTTTTGTAAATTATATTGTGCTATAGTATTGATTCAAAATATTAAAAATG   1657

Db  1648  TCTTGCTGTTGACATATTTAAGTGTTTTAAATGTACAGACATATTTAACTGGTGCACTTTG   1707
          ||   ||||||||||||||  |||||||||||  |||||||| |||||||||||||||||||
Qy  1658  TCTCACTGTTGACATATTTAATGTTTTAAATGTTTACAGACATATTTAACTGGTGCACTTTG   1717

Db  1708  TAAATTCCCTGGGGAAAACTTGCCAGCTAAGGAGGGGGAAAAAATGTTGTT-TCCTAATAT   1766
          ||||| ||||  |     ||  | ||||| || |||   ||| |||| ||| |||   | ||
Qy  1718  TAATTCCCTGAGGTA-CTCGGTAGCTAAGGAGTGGGGAAAAAATGTTGTT-TCCTGACCACAT   1776

Db  1767  CAAATGCAGTA-TAT-TCTTCCGT-CTTTTTA-AGTAATAGAGATTTTTCAGACTGTC   1822
          ||||||| ||| |||  |||| |   |||||||  ||||||||| |||||||||||||
Qy  1777  GTAGTTTATTTCTTTATTCTTTTAACTTAAATAGAGTCTTCAGAGACTTGTCAAAACTATGC   1836

Db  1823  AAGCCTG--AA-AAAATTAAAATGGATGCCTTCGAATAATAAGCAGGATGTTGGCC   1878
          ||| |||   ||||||| || || |||||||||| ||||||||  |||||||||||
Qy  1837  AAGCAAATAAAATAAAATAAAATGAAATCTTGAATAATAAGTAGGATGTTGGTC   1896

Db  1879  ACCAGGTGCCTTCAAATTTAGAAACTAATTGACTTTAGAAAGCTGACATTGCCAAAAAG   1938
          ||||||||||||||||||||||||||||||  |||||||  |||||||||||||||||||
Qy  1897  ACCAGGTGCCTTTCAAATTTGACTTTAGGA-GCTGACATAGGCCAAAAAG   1955

Db  1939  GATACATAATGGGCCACTGAAATT-TGTCAAGAGTAGTATATAATGTTGAACAGGTGT   1997
          ||||||||||||||||||||||  ||||| ||||||||| ||| ||| |||||||||||
Qy  1956  GATACATAATGGCTACTGAAAATCTGTCAGGAGTAGTATTTATGCAATTATTGAACAGGTGT   2015

Db  1998  --TTTTCCACAAGTGCCGCAAATTGTACCTTTT--TT-TTTTTTTCAAAATGAGAAAAGTT   2052
          ||  ||||  |||| || | |  ||||||||  ||  |||||||||  |||||||||||||
Qy  2016  CTTTTTTTACAAGAGCTACAAATTGTAAAATTTTGTTTTTTTTTTTTCCCATAGAAAAATGT   2075

Db  2053  ATTAGTGGTTTATCAGC--AAAAA-AGTCCAATTT-AATTTAGTAAATGTTATTTAT-   2107
          || ||||  |||||||||  ||||| ||  |||||  ||||||||||||||||||||
Qy  2076  ACTA-TAGTTTATCAGCCAAAAAACAATCCACTTTTTAATTTAGTGAAAGTTATTTTATT   2134

Db  2108  --ACTGTACAATAAAACATTGCCTTTGAATGTAATTTTTTGGTACAAAA   2157
          || |||||||| || ||||||  || ||||||||||||||||||||||
Qy  2135  ATACTGTACAATAAAAGCATTGTTTCTGAATGGCA-TTTTTCGTGAATGGCA-TTTTTTCGGTACTTAA   2185
```

```
RESULT  7
LOCUS       AF013171     1823 bp    mRNA           PRI       20-SEP-1997
DEFINITION  Homo sapiens TNF-related ligand TRANCE mRNA, partial cds.
ACCESSION   AF013171
NID         g2411499
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1823)
  AUTHORS   Wong,B.R., Rho,J., Arron,J., Robinson,E., Orlinick,J., Chao,M.,
            Kalachikov,S., Cayani,E., Bartlett,F.S. III, Frankel,W.N., Lee,S.Y.
            and Choi,Y.
  TITLE     TRANCE is a novel ligand of the tumor necrosis factor receptor
            family that activates c-Jun N-terminal kinase in T cells
  JOURNAL   J. Biol. Chem. 272 (40), 25190-25194 (1997)
  MEDLINE   97460112
REFERENCE   2  (bases 1 to 1823)
  AUTHORS   Wong,B.R., Rho,J., Arron,J., Lee,S.Y., Robinson,E. and Choi,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-1997) Howard Hughes Medical Institue, The
            Rockefeller University, 1230 York Ave., New York, NY 10021, USA
FEATURES             Location/Qualifiers
     source          1..1823
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /map="13q14"
                     /cell_type="thymocyte"
     misc_feature    1..738
                     /note="encodes extracellular domain"
```

```
Db    838   GCATGGCCAACG-GTACACGACTCAGTATCCATGCTCTTGACCTTGTAGAGAACACGCG  896
            ||||| |||||| ||||||  || | |||||| ||||||  |||||||||| ||| ||
Qy   1171   ACATGGGCCACGGTGTATGAAACTCACAGCCCTCTCTTCTTGAGCCTGTACAGGTTGTGTA  1230

Db    897   TATTACAGCCAGTGGGAGATGTTAGACTCATGGTGTGTTACACAATGGTTTTAAATTT  956
            ||| ||  ||   ||||  |||||| || || |  |||||||| | |||||  ||||||
Qy   1231   TATGTAAAGTCCATAGGTGATGTGATGATTCATGGTGA-TTACACACGGTTTACAATTT  1289

Db    957   TGTAATGAATTCCTA-GAATTAAACCAGATTGG---AGCAATTACGG-G--T-TGAC---  1005
            |||||| ||||| || |||||||| | ||| || ||||| ||   | |  | ||||
Qy   1290   TGTAATGAGATTCCTAAGAATTGAACCAGATTGGGAGAGGTATTCCGATGCTTATGAAAAA  1349

Db   1006   CTTATGAGAAA-CT--GCATGTGGCTATGGGAG--GGGT-TGGTCCGTGTCATGTGCC  1059
            || |||   || ||  ||| |||||| || |||  || |  | |||| ||||| |||
Qy   1350   CTTACACGTGAGCTATGGGAAGGGGGTCACAGTCTCTGGGTCTAACCCCTGGACATGTGCC  1409

Db   1060   CCTTCGCAGCT-GAAGTGGAGAGGTGTCATCT-AGCGCAATTGAAGGATCATCTGAAGG  1117
            ||   |||| |  | ||||  || | | || || | || |||||||| |||| ||||||
Qy   1410   ACTGAGAACCTTGAAATTAAGAGAGATGCCATGTCATTGCAAAGAAATGATAGTGTGAAGG  1469

Db   1118   GGCAAATTCTTTGAATTGTTACATCATGCTGGAACCTGCAAAAA-TACTTTTT--CTA  1174
            | | |||||||||||||||| ||  || || |   | ||||||| |||   |   ||
Qy   1470   GTTAAGTTCTTTGAATTGTTACATTGCGCTGGGACCTGCAAATAAGTTCTTTTTTTCTA  1529

Db   1175   ATGAGGAGAGAAAA-TATATGTATTTTATATAATAATCTAAAGTTATATTTCAGGTGTAA  1233
            |||||||||||||| |||| |||| |||||| ||| |  |||||||| |||||||| |||
Qy   1530   ATGAGGAGAGAAAAATATATGTATTTTTATATAATGTCTAAAGTTATATTTCAGGTGTAA  1589

Db   1234   TGTTTTCTTTGCAAAGTATTGTAAATTAATTATTGTTGCTATAGTATTTCAAAAATATT  1293
            ||||||||||||||| ||||| |||||| |||||| ||||||||||| ||||||| |||
Qy   1590   TGTTTTCTGTGCAAAGTTTTGTAAATTATTTGTTGCTATAGTATTTCAAAAATATT  1649

Db   1294   TAAAAATGTCTTGCTGTTGACATATTTAAATGTACAGACATATTTAACTGGT  1353
            |||||||||||| ||||||||||||||| |||||| ||||||||||||||||
Qy   1650   TAAAAATGTCTTGACTGTTGACATATTTAAATGTACAGATATTTTAACTGGT  1709

Db   1354   GCACTTTGTAAATTCCCTGGGGAAAACTTGCAGCTAAGGAGGGGAAAAAATGTTGTTCC  1413
            || ||||||||||| |||| | | |||||||| ||| | ||| | | |||||| |||
Qy   1710   GCACTTTGTAATTCCCCTGAAGGTA-CTCGTAGCTAAGGGGCCAGAATACTGTTTTCTGGT  1768

Db   1414   TA--ATATCAAATGCAGTA-TAT-TTCTTCGTT-CTTTTTA-AGTTAATAGATTTTTTCA  1467
            ||  ||| |  |  || |  |||  || || ||| |||  |  ||||||   | ||||||
Qy   1769   GACCACCATGTAGTTTATTTCTTATTCTTTTTTAACTTAATAGAGTCTTCAGACTTGTCAA  1828

Db   1468   GACTTGTCAAGCCTG-TGCA--AA-AAAATTAAAATGGATGCCTTGAATAATAAGCAGGA  1523
            ||||||||||||    |  |  || |||||| |||   ||||  || || |||| || ||
Qy   1829   AACTATGCAAGCAAAATAAATAAAATAAAATAAAATGAATATCTTGAATAATAAGTAGGA  1888

Db   1524   TGTTGGCCACCAGGTGCCTTCAAATTTAGAAACTTGACTTTAGAAAGTCGACATTG  1583
            ||||| || ||||| |||||||| ||||||| ||| |||| |||| | ||||||||
Qy   1889   TGTTGGTCACCAGGTGCCTTTCAAATTTCAAATTGCCAATTGACTTTAGGA-GCTGACATAG  1947

Db   1584   CCAAAAGGATACATAATGGGCCACTGAA-TCTGTCAAGAGTAGTTATATAATTGTTGA  1642
            |||||| ||||| ||| |  |||| |||  |||  |||||| |  |||| ||| |||||
Qy   1948   CCAAAAAGGATACATAATAGGCTACTGAAAATCTGCAGGAGTATTTATGCAATTATTGA  2007

Db   1643   ACAGGTGT--TTTTCCACAGTGCCGCAAATTGTACCTTTT--TTTGTGTTTTTTCAAAATA  1699
            |||||||   ||||||||| | ||||||||||| ||  |||  ||||||||| |||||||
Qy   2008   ACAGGTGTCTTTTTTTTTACAAGAGCTACAAATTGTAAAATTTGTTTCTTTTTTTTTCCCATA  2067

Db   1700   GAAAAGTTATTAGTGGTTTATCAGC--AAAAA-AGTCCAATTT-AATTTAGTAAATGTT  1755
            ||||||||| |  || |||| |||   |||||  ||| ||||  |||||| |||| |||
Qy   2068   GAAAATGTACTA-TAGTTTATCAGCGAAAAAAAACAATCCACTTTTTAATTTAGTAGAAGTT  2126

Db   1756   ATCTTAT---ACTGTACAATAAAAAACATTGCCTTTGAATGTTAATTTTTTGGTAC  1807
            || | ||   |||||||| ||||| |||| | ||||||||| ||| ||||||||||
Qy   2127   ATTTATTATACTGTACAATAAAAAGCATTGTTTCTGAATGGCA-TTTTTTGGTAC  2180
```

RESULT   8

LOCUS   AF012089   10772 bp   DNA   INV   05-AUG-1997

DEFINITION Drosophila melanogaster cysteine proteinase-1 (CP1) gene, complete
           cds, and phenylalanyl tRNA synthetase gene, partial cds.

ACCESSION AF012089
NID   g2305220
KEYWORDS   .
SOURCE   fruit fly.
  ORGANISM Drosophila melanogaster
           Eukaryota; mitochondrial eukaryotes; Metazoa; Arthropoda;
           Tracheata; Hexapoda; Insecta; Pterygota; Diptera; Brachycera;
           Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 4546 to 4553)
  AUTHORS Gray,Y.H., Tanaka,M.M. and Sved,J.A.
  TITLE   P-element-induced recombination in Drosophila melanogaster: hybrid
           element insertion
  JOURNAL Genetics 144 (4), 1601-1610 (1996)
  MEDLINE 97132596
REFERENCE 2 (bases 1 to 10772)
  AUTHORS Gray,Y.H.M., Sved,J.A., Preston,C.R. and Engels,W.R.
  TITLE   structure of the cysteine proteinase (CP1) gene of Drosophila
           melanogaster and associated mutational effects
  JOURNAL Unpublished
REFERENCE 3 (bases 1 to 10772)
  AUTHORS Gray,Y.H.M., Sved,J.A., Preston,C.R. and Engels,W.R.
  TITLE   Direct Submission
  JOURNAL Submitted (30-JUN-1997) School of Biological Sciences, University
           of Sydney, Biology A12, Sydney University, NSW 2006, Australia
FEATURES   Location/Qualifiers
     source   1..10772
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
     mRNA   join(872..1000,2310..2426,6476..6690,6751..7707)
                     /gene="CP1"
                     /product="cysteine protease"
     gene   872..7707
                     /gene="CP1"
     exon   872..1000
                     /gene="CP1"
                     /number=1
     intron   1001..2309
                     /gene="CP1"
     exon   2310..2426
                     /gene="CP1"
                     /number=2
     CDS   join(2328..2426,6476..6690,6751..7462)
                     /gene="CP1"
                     /codon_start=1
                     /product="cysteine proteinase-1"
                     /db_xref="PID:g2305221"
                     /translation="MRTAVLLPLLALLAVAQAVSFADVVMEEWHTFKLEHRKNYQDET
                     EERFRLKIFENEKHKIAKHNQRFAEGKVSFKLAVNKYADLLHEFRQLMNGFNYTLHK
                     QLRAADESFKGVTFISPAHVTLPKSVDWRTKGAVTAVKDQGHCGSCWAFSSTGALEGQ
                     HFRKSGVLVSLSEQNLVDCSTKYGNNGCNGGLMDNAFRYIKDNGGIDTEKSYPYEAID
                     DSCHFNKGTVGATDRGFTDIPQGDEKKMAEAVATVGPVSVAIDASHESFQFYSEGVYN
                     EPQCDAQNLDHGVLVVGFGTDESGEDYWLVKNSWGTTWGDKGFIKMLRNKENQGIAS
                     ASYPIV"
     intron   2427..6475
                     /gene="CP1"
     misc_feature   4546..4553
                     /gene="CP1"
                     /note="insertion site of P[CaSpeR](50C)"
                     /citation=[1]
     exon   6476..6690
                     /gene="CP1"
                     /number=3
     intron   6691..6750
                     /gene="CP1"
     exon   6751..7707
                     /gene="CP1"
                     /number=4
     mRNA   join(<8110..9300,9370..>9532)
                     /product="phenylalanyl tRNA synthetase"
     CDS   join(8110..9300,9370..>9532)
                     /note="potential orf"
                     /codon_start=1
                     /product="phenylalanyl tRNA synthetase"

```
            /db_xref="PID:g2305222"
            /translation="MLLTLRVQGARHWLKSTRCLASSAAPAKSPSSPPQLEVSGSTYA
            TDGWTNVTPKILSYVGANKHLQTDHPLSIIRQRIVNYFYGAYRNQRGNPLFSVYDQMN
            PVVTVQQNFDNLLIPADHVSRQKSDCYYINQQHLLRAHTAHQVELISGGLDNFLVVG
            EVYRRDEIDSTHYPVFHQADAVRLVTKDKLFERNPGLELFEETWSGTLADPKLILPHP
            SSWTKPNSPATRRAVKLMEHEMKHVLGLTKDLFGPRIKYRWVDTYFPFTQPSWELEI
            YFKDNWLEVLGCGIMRHEILQRSGVHQSIGYAFGVGLERLAMVLFDIPDIRLFWSNDS
            GFLSQFSEKDLHNLPKYKPISHYPQCTNDLSFWLPQDIEVDAGFSPNDFVDLVRSVAG
            DMVEQISLVDKFKHPKTGKSSVCFRIVYRHMERTLTQAEVNEIHKQIASASVDSFNVQ
            IX"
BASE COUNT     2929 a   2357 c   2282 g   3046 t    158 others
ORIGIN

Query Match          2.5%;    Score 55;   DB 14;  Length 10772;
Best Local Similarity 18.4%;    Pred. No. 6.20e-15;
Matches  27;  Conservative  79;  Mismatches  39;  Indels  2;  Gaps  2;

Db   1687  KTAAAKRWWYWAWTTTMWKMWTTWKWAMMKTYRTWMWKM-YWTSRTTTTSAMWMWYTW  1745
            :||  :|::::|  :||:||  :|: ||: |: :|: ::: |:  ||::|:|:||::|:
Qy   1989  GTATTTATGCAATTATTGAACAGGTGTCTTTTTTTTACAAGAGCTACAAAATTGTAAATTTT  2048

Db   1746  STWT-KYWWAYAWMKMWWTRTWARMAWASWARWKWKTSAAAAYSAWRKMWKWWAYRAMKK  1804
            :|:| :|:::|  ::::: :|::: |::||: ::: :|: ::|::  |::|:
Qy   2049  GTTCTTTTTTTTTCCCATAGAAAATGTACTATAGTTTATCAGCCAAAAAACAATCCACTT  2108

Db   1805  TWMWAAWKRWRKAAWTWRWWYMTTA  1831
            |: | ::::::|| |:::::| |||
Qy   2109  TTTAATTTAGTGAAAGTATTTTATTA  2135


RESULT   9
LOCUS       AF012089   10772 bp    DNA             INV      05-AUG-1997
DEFINITION  Drosophila melanogaster cysteine proteinase-1 (CP1) gene, complete
            cds, and phenylalanyl tRNA synthetase gene, partial cds.
ACCESSION   AF012089
NID         g2305220
KEYWORDS    .
SOURCE      fruit fly.
ORGANISM    Drosophila melanogaster
            Eukaryota; mitochondrial eukaryotes; Metazoa; Arthropoda;
            Tracheata; Hexapoda; Insecta; Pterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 4546 to 4553)
AUTHORS     Gray,Y.H., Tanaka,M.M. and Sved,J.A.
TITLE       P-element-induced recombination in Drosophila melanogaster: hybrid
            element insertion
JOURNAL     Genetics 144 (4), 1601-1610 (1996)
MEDLINE     97132596
REFERENCE   2  (bases 1 to 10772)
AUTHORS     Gray,Y.H.M., Sved,J.A., Preston,C.R. and Engels,W.R.
TITLE       Structure of the cysteine proteinase (CP1) gene of Drosophila
            melanogaster and associated mutational effects
JOURNAL     Unpublished
REFERENCE   3  (bases 1 to 10772)
AUTHORS     Gray,Y.H.M., Sved,J.A., Preston,C.R. and Engels,W.R.
TITLE       Direct Submission
JOURNAL     Submitted (30-JUN-1997) School of Biological Sciences, University
            of Sydney, Biology A12, Sydney University, NSW 2006, Australia
FEATURES    Location/Qualifiers
source      1..10772
            /organism="Drosophila melanogaster"
            /db_xref="taxon:7227"
mRNA        join(872..1000,2310..2426,6476..6690,6751..7707)
            /gene="CP1"
            /product="cysteine protease"
gene        872..7707
            /gene="CP1"
exon        872..1000
            /gene="CP1"
            /number=1
intron      1001..2309
            /gene="CP1"
exon        2310..2426
            /gene="CP1"
```

```
CDS         join(2328..2426,6476..6690,6751..7462)
            /gene="CP1"
            /number=2
            /codon_start=1
            /product="cysteine proteinase-1"
            /db_xref="PID:g2305221"
            /translation="MRTAVLLPLLALLAVAQAVSFADVVMEEWHTFKLEHRKNYQDET
            EERFRLKIFNENKHKIAKHNQRFAEGKVSFKLAVNKYADLLHHEFRQLMNGFNYTLHK
            QLRAADESFKGVTFISPAHVTLPKSVDWRTKGAVTAVKDQGHCGSCWAFSSTGALEGQ
            HFRKSGVLVSLSEQNLVDCSTKYCNNGCNGGLMDNAFRYIKDNGGIDTEKSYPYEAID
            DSCHFNKGTVGATDRGFTDIPQGDEKKMAEAVATVGPVSVAIDASHESFQFYSEGVYN
            EPQCDAQNLDHGVLVVGFGTDESGEDYWLVKNSWGTTWGDKGFIKMLRNKENQCGIAS
            ASSYPLV"
intron          2427..6475
                /gene="CP1"
                4546..4553
misc_feature    /gene="CP1"
                /note="insertion site of P[CaSpeR](50C)"
                6476..6690
exon            /citation=[1]
                /gene="CP1"
                /number=3
                6691..6750
intron          /gene="CP1"
                6751..7707
exon            /gene="CP1"
                /number=4
mRNA            join(<8110..9300,9370..>9532)
                /product="phenylalanyl tRNA synthetase"
                join(8110..9300,9370..>9532)
CDS             /note="potential orf"
                /codon_start=1
                /product="phenylalanyl tRNA synthetase"
                /db_xref="PID:g2305222"
                /translation="MLLTLRVQGARHWLKSTRCLASSAAPAKSPSSPPQLEVSGSTYA
            TDGWTNVTPKILSYVGANKHLQTDHPLSIIRQRIVNYFYGAYRNQRGNPLFSVYDQMN
            PVVTVQQNFDNLLIPADHVSRQKSDCYYINQQHLLRAHTAHQVELISGGLDNFLVVG
            EVYRRDEIDSTHYPVFHQADAVRLVTKDKLFERNPGLELFEETWSGTLADPKLILPHP
            SSWTKPNSPATRRAVKLMEHEMKHVLGLTKDLFGPRIKYRWVDTYFPFTQPSWELEI
            YFKDNWLEVLGCGIMRHEILQRSGVHQSIGYAFGVGLERLAMVLFDIPDIRLFWSNDS
            GFLSQFSEKDLHNLPKYKPISHYPQCTNDLSFWLPQDIEVDAGFSPNDFVDLVRSVAG
            DMVEQISLVDKFKHPKTGKSSVCFRIVYRHMERTLTQAEVNEIHKQIASASVDSFNVQ
            IX"
BASE COUNT     2929 a   2357 c   2282 g   3046 t    158 others
ORIGIN

Query Match              ;    Score 52;   DB 14;  Length 10772;
Best Local Similarity 18.3%;    Pred. No. 3.84e-13;
Matches  26;  Conservative  76;  Mismatches  38;  Indels  2;  Gaps  2;

Db   1688  TAAAKRWWYWAWTTTMWKMWWTTWKWAMMKTYRTWMWKMYWTSRTTTTSAMWMWYTWST  1747
            |||  :::::|||:||| |: : ::::| |: ::::|::|  :|:::| |:||
Cp   2135  TAATAAAATAACTTCACTAAATTAAAAAGTGGATTGTTTTTGGCTGATAAACTATAGT  2076

Db   1748  WT-KYWWAYAWMKMWWTRTWARMAWASWARWKWKTSAAAAYSAWRKMWKWWAYRAMKKTW  1806
            :|:::||: |||::: :: :|: ||::::::  :||::|| :||
Cp   2075  ACATTTCTATGGAAAAAAAGAACAAATTTACAATTG-TAGCTCTTGTAAAAAAA  2017

Db   1807  MWAAWKWRWKAAWTWRWWYMT  1828
            |: :::::| :|:::::|:|
Cp   2016  GACACCGTTCAATAATTGCAT  1995


RESULT   10
LOCUS       I66494      7218 bp    DNA           PAT      23-DEC-1997
DEFINITION  Sequence 14 from patent US 5670367.
ACCESSION   I66494
NID         g2724471
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
```

REFERENCE   1 (bases 1 to 7218)
AUTHORS   Dorner,F., Scheiflinger,F. and Falkner,F.Gunter.
TITLE   Recombinant fowlpox virus
JOURNAL   Patent: US 5670367-A 14 23-SEP-1997;
FEATURES   Location/Qualifiers
source   1..7218
            /organism="unknown"
BASE COUNT 1944 a   1491 c   1486 g   1929 t   368 others
ORIGIN

Query Match   2.1%; Score 47; DB 17; Length 7218;
Best Local Similarity 7.9%; Pred. No. 3.22e-10;
Matches   14; Conservative 98; Mismatches 65; Indels 0; Gaps 0;

Db   1045 GCAGGTCGAGGGAGCTTCGCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1104
           |||| ||| ||||-||| ||-|| :: :::: ::: :: : :::::
Qy   162 GCAGCTCGGAGGAGATGGGCAGCCGGGTCCCCACAGAGGGTCCGCTGCACCCCG 221

Db   1105 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1164
           :::::::::::: :: :::: :::::::::::::::::::::::::::::
Qy   222 CGCCTCTCGCACCGGCTCCGGCGGCCGCGGGCCGCTCCCCGCCTCCATGTTCTCGGCCC 281

Db   1165 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1221
           :: :: ::::: :: ::::::::::::::::::::::: : : :
Qy   282 TCCTGGGGCTGGGACTGGGCCAGGTGGTCTGCAGCATCGCTGCTCGTTCCTGTACTTTC 338

RESULT 11
LOCUS   I66494   7218 bp   DNA   PAT   23-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION   I66494
NID   g2724471
KEYWORDS   .
SOURCE   Unknown.
ORGANISM   Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 7218)
AUTHORS   Dorner,F., Scheiflinger,F. and Falkner,F.Gunter.
TITLE   Recombinant fowlpox virus
JOURNAL   Patent: US 5670367-A 14 23-SEP-1997;
FEATURES   Location/Qualifiers
source   1..7218
            /organism="unknown"
BASE COUNT 1944 a   1491 c   1486 g   1929 t   368 others
ORIGIN

Query Match   1.9%; Score 41; DB 17; Length 7218;
Best Local Similarity 0.5%; Pred. No. 7.68e-07;
Matches   1; Conservative 113; Mismatches 73; Indels 0; Gaps 0;

Db   1249 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1308
           ::::::: :: :::::::: :::::::::: ::::: :::::::::
Cp   2012 CCTGTTCAATAATTGCATAAAATACTCCTGACAGATTTTCAGTAGCCTATTATGTATCCTT 1953

Db   1309 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1368
           :::: ::: :: :::::::: :: :: :::::: :::: :
Cp   1952 TTTGGCTATGTCAGCTCCTAAAGTCAATTAGCTTCTAAATTTGAAAGGCACCTGGTGACC 1893

Db   1369 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1428
           :: ::::: ::: :: ::: :: ::::::: :: ::: :::::: :
Cp   1892 AACATCCTACTTATTATTCAAGATATTCATTTTTATTTTTATTTTATTTTGCTTGCAT 1833

Db   1429 YYYYYG 1435
           ::::: |
Cp   1832 AGTTTTG 1826

RESULT 12
LOCUS   I28278   215 bp   DNA   PAT   30-OCT-1996
DEFINITION Sequence 5 from patent US 5569830.
ACCESSION   I28278
NID   g1819054
KEYWORDS   .

SOURCE   Unknown.
ORGANISM   Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 215)
AUTHORS   Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.
TITLE   Plant inhibitors of fungal polygalacturonases and their use to
            control fungal disease
JOURNAL   Patent: US 5569830-A 5 29-OCT-1996;
FEATURES   Location/Qualifiers
source   1..215
            /organism="unknown"
BASE COUNT   15 a   8 c   25 g   26 t   141 others
ORIGIN

Query Match   1.8%; Score 39; DB 17; Length 215;
Best Local Similarity 12.5%; Pred. No. 9.41e-06;
Matches   26; Conservative 85; Mismatches 96; Indels 1; Gaps 1;

Db   8 SSSVVSRTASCNDKAKKDGNTTSSWTTDCCNRTWGVCDTDTTYRVNNDSGHNKYSSANYN 67
           ::: ::: :: :: : |: || :::| : :: :::
Cp   385 GCAGYGAGTCGTGTTCTCGATAATTCTGTTAGGATCCATCTGCGCTCGAAAGTACAGGAA 326

Db   68 YGGNNVGAAKTHYTHTNVSGADSKTVTDSYNASGTSSSNGGTDGNRSGADSYGSSKTAM 127
           :::: :: :: ::| | : :: |:: ::: ::: :
Cp   325 CAGAGCGATGCTGCAGACCACCTGGCCCAGTCCCAGTCCCAGGAGGGCCAGGAAACATGGA 266

Db   128 TSRNRTGKTANNAVDSRNMGDASV-GSDKNTKKHAKNSADGKVGSKNNGDRNNRYGTGTK 186
           :::: :: ::: : :: ::::: |::::: :::: :: :::|
Cp   265 GCGGGAGGCGGCCGGCCGGCGGGTGGCCGGCCGGAGCCGGTGCAGAAGGCGCGGGGTGCAGCGGACC 206

Db   187 SNVSNNCGGNKRDVSSYANNKCCGSSC 214
           ||| ||| :::| : ::::|
Cp   205 CTCGTGTGGGACGGCGGGGCCGCTGCCC 178

RESULT 13
LOCUS   AC005369   74371 bp   DNA   PRI   01-AUG-1998
DEFINITION Homo sapiens chromosome 5, BAC clone 119J3 (LBNL H175), complete
            sequence.
ACCESSION   AC005369
NID   g3367505
KEYWORDS   HTG.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 74371)
AUTHORS   Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
            Kadner,K., Miguel,T., Miller,C., Pitluck,S., Pollard,M.,
            Rojeski,H., Subramanian,S. and Martin,C.H.
TITLE   Sequencing of human chromosome 5
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 74371)
AUTHORS   Ricke,D.O.
TITLE   Large Scale Sequence Analysis and Annotation with the Sequence
            Comparison Analysis (SCAN) System
JOURNAL   Unpublished
REFERENCE   3 (bases 1 to 74371)
AUTHORS   Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
            Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M.,
            Rojeski,H., Subramanian,S. and Martin,C.H.
TITLE   Direct Submission
JOURNAL   Submitted (01-AUG-1998) Human Genome Center, DOE Joint Genome
            Institute, Lawrence Berkeley National Laboratory, MS 74-157,
            Berkeley, CA 94720, U.S.A.
COMMENT   Sequence submitted by:
            DOE Joint Genome Institute.
FEATURES   Location/Qualifiers
source   1..74371
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /map="5q"
            /clone="119J3"

```
repeat_region    /chromosome="5"
                 /note="LBNL H175"
                 893. .1030
repeat_region    /rpt_family="Alu"
                 2295. .2438
repeat_region    /rpt_family="Alu"
                 2818. .2859
                 /note="(GT)21"
                 /rpt_type=tandem
                 /rpt_unit=GT
misc_feature     join(3246. .3410,3721. .3828)
                 /standard_name="RLF"
                 /note="65% & 69% protein identity GenPept:U22377"
repeat_region    3431. .3724
                 /rpt_family="Alu"
repeat_region    3707. .3728
                 /note="(A)22"
                 /rpt_type=tandem
                 /rpt_unit=A
repeat_region    4366. .4661
                 /rpt_family="Alu"
repeat_region    5327. .5602
                 /rpt_family="Alu"
repeat_region    6586. .6956
                 /rpt_family="L1"
repeat_region    6647. .6684
                 /note="(CA)19"
                 /rpt_type=tandem
                 /rpt_unit=CA
repeat_region    7113. .7373
                 /rpt_family="Alu"
repeat_region    complement(7830. .8185)
                 /standard_name="possible repeat"
repeat_region    8258. .8503
                 /rpt_family="Alu"
repeat_region    9070. .9387
                 /rpt_family="Alu"
repeat_region    complement(9740. .9845)
                 /rpt_family="MER42"
repeat_region    complement(10440. .11015)
                 /rpt_family="Alu"
repeat_region    11950. .12250
                 /rpt_family="Alu"
repeat_region    12057. .12085
                 /note="(A)29"
                 /rpt_type=tandem
                 /rpt_unit=A
repeat_region    12365. .12645
                 /rpt_family="Alu"
repeat_region    13727. .13750
                 /note="(AC)12"
                 /rpt_type=tandem
                 /rpt_unit=AC
repeat_region    13783. .14024
                 /rpt_family="L1"
repeat_region    14175. .14470
                 /rpt_family="Alu"
repeat_region    complement(14906. .15259)
                 /standard_name="possible repeat"
repeat_region    15300. .15613
                 /rpt_family="Alu"
repeat_region    16671. .16690
                 /note="(T)20"
                 /rpt_type=tandem
                 /rpt_unit=T
repeat_region    complement(16675. .16977)
                 /rpt_family="Alu"
repeat_region    complement(16993. .17085)
                 /rpt_family="MER42"
repeat_region    complement(17678. .18276)
                 /rpt_family="Alu"
repeat_region    19305. .19583
                 /rpt_family="Alu"

repeat_region    19914. .19945
                 /note="(GTTT)8"
                 /rpt_type=tandem
                 /rpt_unit=GTTT
repeat_region    complement(19943. .20222)
                 /rpt_family="Alu"
repeat_region    21202. .21496
                 /rpt_family="Alu"
repeat_region    21736. .22035
                 /rpt_family="Alu"
repeat_region    22017. .22038
                 /note="(A)22"
                 /rpt_type=tandem
                 /rpt_unit=A
repeat_region    22321. .22457
                 /rpt_family="Alu"
repeat_region    complement(22638. .22981)
                 /standard_name="possible repeat"
repeat_region    23022. .23326
                 /rpt_family="Alu"
repeat_region    23473. .23761
                 /rpt_family="Alu"
repeat_region    23744. .23767
                 /note="(A)24"
                 /rpt_type=tandem
                 /rpt_unit=A
repeat_region    complement(24639. .24694)
                 /rpt_family="MER42"
repeat_region    complement(25349. .25713)
                 /rpt_family="Alu"
repeat_region    complement(25727. .26471)
                 /rpt_family="Alu"
repeat_region    27191. .27477
                 /rpt_family="Alu"
repeat_region    27774. .28057
                 /rpt_family="Alu"
repeat_region    28040. .28066
                 /note="(A)27"
                 /rpt_type=tandem
                 /rpt_unit=A
repeat_region    28712. .28930
                 /rpt_family="MER20"
misc_feature     complement(28769. .28838)
                 /note="GRAIL 2 excellent exon, frame 0"
repeat_region    28987. .29214
                 /rpt_family="Alu"
misc_feature     join(29485. .29559,29598. .29779)
                 /note="95% identity dbSTS:G14522 (SHGC-11312)"
misc_feature     29495. .29976
                 /standard_name="AI027942"
                 /note="100% identity EST ov84a10.x1"
misc_feature     complement(30401. .30536)
                 /note="GRAIL 2 excellent exon, frame 2"
repeat_region    complement(30682. .30733)
                 /rpt_family="MIR"
misc_feature     complement(31573. .31724)
                 /note="GRAIL 2 excellent exon, frame 1"
misc_feature     complement(32159. .32232)
                 /note="GRAIL 2 excellent exon, frame 2"
repeat_region    32388. .32488
                 /rpt_family="MLT1"
repeat_region    32617. .32908
                 /rpt_family="Alu"
repeat_region    32977. .33088
                 /rpt_family="MLT1"
repeat_region    complement(33670. .33785)
                 /rpt_family="Alu"
misc_feature     complement(34021. .34144)
                 /note="GRAIL 2 excellent exon, frame 2"
misc_feature     complement(35238. .35331)
                 /note="GRAIL 2 excellent exon, frame 1"
repeat_region    36392. .36663
                 /rpt_family="Alu"
```

```
misc_feature        36901..37222
                    /note="100% identity EST ou55c09.x1"
                    /db_xref="dbEST:AI025011"
STS                 36901..37164
                    /standard_name="STSG-9983"
                    /db_xref="dbSTS:G26554"
misc_feature        complement(37595..37654)
                    /note="GRAIL 2 excellent exon, frame 0"
misc_feature        complement(join(38063..38218,38462..38578,38741..38995,
                    39071..39205,39532..39630,39935..40048,40300..40410,
                    40503..40661,41868..41972,42103..42225,42492..42569,
                    44379..44507))
                    /standard_name="histidyl-tRNA synthetase"
                    /note="78%-100% protein identity GenPept:U18937"
misc_feature        complement(38069..38215)
                    /note="GRAIL 2 excellent exon, frame 0"
misc_feature        complement(38462..38578)
```

Note: remainder of annotations omitted.

```
Query Match         1.8%;     Score 40;  DB 21;  Length 74371;
Best Local Similarity 11.7%;  Pred. No. 2.70e-06;
Matches     26;   Conservative  108;  Mismatches  88;  Indels   1;  Gaps   1;

Db  15931  SRSKSKRGWGYRSWKKYRCAMWMTCKSSKCW-CWSYRMRMKCYSCSYCYCSSGKKYWCRC 15989
                ::::::  :: ::  :::::::! !::::  !:!::::!  :::::::::!:  :!!  :!
Qy     89  CGCGGAGCAGGCGCGCCGAACTCCGGCGCCGCCCATGCCGGGCCCAGCCGAGACTAC 148

Db  15990  SMYWTCYYSYKYYWSMSYCTCTSWGWRMWSKGRSWMYASRSGCSCSCSMCMMCRCSCM 16049
                :!: :: ::  ::   ::! :::: :  :::!  :!: ::::::::::::::::::::
Qy    149  GGCAAGTACCTGCCGAGCTCGGAGAGAGATGGCAGCGGCCCCGGCGTCCCACACGAGGGT 208

Db  16050  SMKMWWTTTTTKTRTWRTTTWKWRKAGASASRGKSKCRCMSYGKKSTCKMKMTCYYGMYS 16109
                :::::::       :: ::!!! :: ::::!!  ::::  ::::: !!:::  ::::::
Qy    209  CCGCTGCACCCCGGCCTTCTGCACCGGCTCCGGCGCCGACACCCGGCGCCTCCCGCTCC 268

Db  16110  WCSYKMKYSRYCCGCCYCSSSYCKSCCWMMSWRMKKGSWKKRMW 16152
                !:::::::  ::::!! !!:::::  ! :::::::  ! !!:::  ::
Qy    269  ATGTTCCTGGCCCTCCTGGGGCGTGGGACTGGGCCAGGTGGTCT 311
```

```
RESULT   14
LOCUS     AC005369     74371 bp     DNA              PRI    01-AUG-1998
DEFINITION Homo sapiens chromosome 5, BAC clone 11j3 (LBNL H175), complete
          sequence.
ACCESSION AC005369
NID       g3367505
KEYWORDS  HTG.
SOURCE    human.
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
          Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1  (bases 1 to 74371)
AUTHORS   Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
          Kadner,K., Miguel,T., Miller,C., Pitluck,S., Pollard,M.,
          Rojeski,H., Subramanian,S. and Martin,C.H.
TITLE     Sequencing of human chromosome 5
JOURNAL   Unpublished
REFERENCE 2  (bases 1 to 74371)
AUTHORS   Ricke,D.O.
TITLE     Direct Submission
JOURNAL   Submitted (01-AUG-1998) Human Genome Center, DOE Joint Genome
          Institute, Lawrence Berkeley National Laboratory, MS 74-157,
          Berkeley, CA 94720, U.S.A.
          Large Scale Sequence Analysis and Annotation with the Sequence
          Comparison Analysis (SCAN) System
REFERENCE 3  (bases 1 to 74371)
AUTHORS   Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
          Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M.,
          Rojeski,H., Subramanian,S. and Martin,C.H.
TITLE     Direct Submission
JOURNAL   Sequence submitted by:
          DOE Joint Genome Institute.
COMMENT
```

```
FEATURES             Location/Qualifiers
source               1..74371
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="5q"
                     /clone="11j3"
                     /chromosome="5"
                     /note="LBNL H175"
repeat_region        893..1030
                     /rpt_family="Alu"
repeat_region        2295..2438
                     /rpt_family="Alu"
repeat_region        2818..2859
                     /note="(GT)21"
                     /rpt_type=tandem
                     /rpt_unit=GT
misc_feature         join(3246..3410,3721..3828)
                     /standard_name="RLF"
                     /note="55% & 69% protein identity GenPept:U22377"
repeat_region        3431..3724
                     /rpt_family="Alu"
repeat_region        3707..3728
                     /note="(A)22"
                     /rpt_type=tandem
                     /rpt_unit=A
repeat_region        4366..4661
                     /rpt_family="Alu"
repeat_region        5327..5602
                     /rpt_family="Alu"
repeat_region        6586..6956
                     /rpt_family="L1"
repeat_region        6647..6684
                     /note="(CA)19"
                     /rpt_type=tandem
                     /rpt_unit=CA
repeat_region        7113..7373
                     /rpt_family="Alu"
repeat_region        complement(7830..8185)
                     /standard_name="possible repeat"
repeat_region        8258..8503
                     /rpt_family="Alu"
repeat_region        9070..9387
                     /rpt_family="Alu"
repeat_region        complement(9740..9845)
                     /rpt_family="MER42"
repeat_region        complement(10040..11015)
                     /rpt_family="Alu"
repeat_region        11950..12250
                     /rpt_family="Alu"
repeat_region        12057..12085
                     /note="(A)29"
                     /rpt_type=tandem
                     /rpt_unit=A
repeat_region        12365..12645
                     /rpt_family="Alu"
repeat_region        13727..13750
                     /note="(AC)12"
                     /rpt_type=tandem
                     /rpt_unit=AC
repeat_region        13783..14024
                     /rpt_family="L1"
repeat_region        14175..14470
                     /rpt_family="Alu"
repeat_region        complement(14906..15259)
                     /standard_name="possible repeat"
repeat_region        15300..15613
                     /rpt_family="Alu"
repeat_region        16671..16690
                     /note="(T)20"
                     /rpt_type=tandem
                     /rpt_unit=T
repeat_region        complement(16675..16977)
                     /rpt_family="Alu"
```

```
repeat_region    complement(16993..17085)
                 /rpt_family="MER42"
repeat_region    complement(17678..18276)
                 /rpt_family="Alu"
repeat_region    19305..19583
                 /rpt_family="Alu"
repeat_region    19914..19945
                 /note="(GTTT)8"
                 /rpt_type=tandem
                 /rpt_unit=GTTT
repeat_region    complement(19943..20222)
                 /rpt_family="Alu"
repeat_region    21202..21496
                 /rpt_family="Alu"
repeat_region    21736..22035
                 /rpt_family="Alu"
repeat_region    22017..22038
                 /note="(A)22"
                 /rpt_type=tandem
                 /rpt_unit=A
repeat_region    22321..22457
                 /rpt_family="Alu"
repeat_region    complement(22638..22981)
                 /standard_name="possible repeat"
repeat_region    23022..23326
                 /rpt_family="Alu"
repeat_region    23473..23761
                 /rpt_family="Alu"
repeat_region    23744..23767
                 /note="(A)24"
                 /rpt_type=tandem
                 /rpt_unit=A
repeat_region    complement(24639..24694)
                 /rpt_family="MER42"
repeat_region    25349..25713
                 /rpt_family="Alu"
repeat_region    complement(25727..26471)
                 /rpt_family="Alu"
repeat_region    27191..27477
                 /rpt_family="Alu"
repeat_region    27774..28057
                 /rpt_family="Alu"
repeat_region    28040..28066
                 /note="(A)27"
                 /rpt_type=tandem
                 /rpt_unit=A
repeat_region    28712..28930
                 /rpt_family="MER20"
misc_feature     complement(28769..28838)
                 /note="GRAIL 2 excellent exon, frame 0"
repeat_region    28987..29214
                 /rpt_family="Alu"
misc_feature     join(29485..29559,29598..29779)
                 /note="95% identity dbSTS:G14522 (SHGC-11312)"
misc_feature     29495..29976
                 /standard_name="AI027942"
                 /note="100% identity EST ov84a10.x1"
misc_feature     complement(30401..30536)
                 /note="GRAIL 2 excellent exon, frame 2"
repeat_region    complement(30682..30733)
                 /rpt_family="MIR"
misc_feature     complement(31573..31724)
                 /note="GRAIL 2 excellent exon, frame 1"
misc_feature     complement(32159..32232)
                 /note="GRAIL 2 excellent exon, frame 2"
repeat_region    32388..32488
                 /rpt_family="MLT1"
repeat_region    32617..32908
                 /rpt_family="Alu"
repeat_region    32977..33088
                 /rpt_family="MLT1"
repeat_region    complement(33670..33785)
                 /rpt_family="Alu"
```

```
misc_feature     complement(34021..34144)
                 /note="GRAIL 2 excellent exon, frame 2"
misc_feature     complement(35238..35331)
                 /note="GRAIL 2 excellent exon, frame 1"
repeat_region    36392..36663
                 /rpt_family="Alu"
misc_feature     36901..37222
                 /note="100% identity EST ou55c09.x1"
                 /db_xref="dbEST:AI025011"
STS              36901..37164
                 /standard_name="STSG-9983"
                 /db_xref="dbSTS:G26554"
misc_feature     complement(37595..37654)
                 /note="GRAIL 2 excellent exon, frame 0"
misc_feature     complement(join(38063..38218,38462..38578,38741..38995,
                 39071..39205,39532..39630,39935..40048,40300..40410,
                 40503..40661,41868..41972,42103..42225,42492..42569,
                 44379..44507))
                 /standard_name="histidyl-tRNA synthetase"
                 /note="78%-100% protein identity GenPept:U18937"
misc_feature     complement(38069..38215)
misc_feature     complement(38462..38578)
                 /note="GRAIL 2 excellent exon, frame 0"
```

Note: remainder of annotations omitted.

```
Query Match            1.7%;  Score 37;  DB 21;  Length 74371;
Best Local Similarity 18.0%;  Pred. No. 1.10e-04;
Matches 24;  Conservative 66;  Mismatches 41;  Indels  2;  Gaps  2;

Db  15923 STSKSCCMSRSKSKRGWGYRSWKKYRCAMWMTCKSSKCWCWSYRMRMKCYSSCSYCYCSSG 15982
            ::|::|: :: :|| :| | |:: :::::|| | :|| :: |::| | | :|:| :: ||
Cp    145 GTCTCGGCTGGCCGGGCGCA-TGGCGCGGCCCGGAGTTCGGGGCCCCTGCTCCGCCGGAT 87

Db  15983 KKYWCRCSMYWYTCYYSYKYYWSMS-YCTCTSWGWRWMWSKGRSWMYASRSGCSCSCSMC 16041
            :|: ::| :|: : :|| |:| : | :|:|: | :| | | :|| : | | :::|
Cp     86 CGTTCTCTCCCTTCCTCCGGCCACCCCGAGCAGACTCCCGCCCAGGCGGCCCCCGC 27

Db  16042 MMCRCSCSMSMKMW 16054
            ::|:::| ::|::
Cp     26 CCCGACCGGTTCA 14
```

```
RESULT   15
LOCUS       I28278                   215 bp    DNA            PAT       30-OCT-1996
DEFINITION  Sequence 5 from patent US 5569830.
ACCESSION   I28278
NID         g1819054
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 215)
AUTHORS     Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.
TITLE       Plant inhibitors of fungal polygalacturonases and their use to
            control fungal disease
JOURNAL     Patent: US 5569830-A 5 29-OCT-1996;
FEATURES             Location/Qualifiers
    source           1..215
                     /organism="unknown"
BASE COUNT       15 a      8 c     25 g     26 t    141 others
ORIGIN
                 1..215
```

```
Query Match            1.6%;  Score 35;  DB 17;  Length 215;
Best Local Similarity 15.9%;  Pred. No. 1.21e-03;
Matches 33;  Conservative 83;  Mismatches 87;  Indels  4;  Gaps  4;

Db  11 VVSRTASCNDKAKKDGNTTSSWTTDCCNRTWGVCDTDTTYRVNNDSG-HNKYSSANYNYG 69
          ::|: | | ::|: ::| :|| | : :: ||  | : :::| |::
Qy  17 ACCGGTCGGGCGGCGGGCCCTGGCCCGGCCGGAGTCTGCTCGGCGGGTGGGTGGCCGAGGAAG 76

Db  70 GNNV-GAAKTHYTHTNVSGADSKTVTDSYNASGTSSSNGGTDGNRSGADSYGSSKTAMT 128
```

```
Qy    77 GGAGAGAACGATCGGCGGAGCAGGGCGCCCGAACTCCGGGCGCCCGCCCATGCGCCGGGC-  135

Db   129 SRNRTGKTANNAVDSRNMGDASVGSDKNTK-KHAKNSADGKVGSKNNGDRNNRYGTGTKS  187
         ::  -| -|::: :|:|:::  :  :  :|  :  :| | |:: |-|:  :|  - : -:
Qy   136 CAGCCGAGACTACGGCAAGTACCTGCGCAGCTCGGAGGAGATGGGCAGCGCGCCCCGGCGT  195

Db   188 NVSNNCGGGNKRDVSSYANNKCCGSSC  214
         ::   |  :::::::::  || :::|
Qy   196 CCCACACGAGGGTCCCGCTGCACCCCGC  222
```

Search completed: Sat Nov 14 02:08:39 1998
Job time : 5347 secs.